The Electronic Library

Early warning of risks of copyright infringement in digital library based on extension theory

Chan Li Wen-De Zhang Yi-Xin Lan

Received 16 April 2014
Revised 26 March 2015
Accepted 13 April 2015

Early warning of risks of copyright infringement in digital library based on extension theory

Chan Li
College of Public Administration, Fujian Agriculture and Forestry University, Fuzhou, China

Wen-De Zhang
Institute of Information Management, Fuzhou University, Fuzhou, China, and

Yi-Xin Lan
School of Economics and Management, Fuzhou University, Fuzhou, China

Abstract

Purpose – The purpose of this study is to evaluate the potential risks of copyright infringement in digital library based on the extension theory.

Design/methodology/approach – At first, the analytic hierarchy process (AHP) is used to determine the weights of the existing indicator system for early warning. Second, a model is built to evaluate the potential risks of copyright infringement based on the extension theory in digital library. Finally, a real-world application is presented to show the effectiveness and usefulness of this approach.

Findings – The main findings of this paper are as follows: the early warning extension theory model is effective in distinguishing the degree of the potential risks of copyright infringement in digital library; the ranges of the value and the values of the indicators can directly affect the results while using this approach, so the accuracy of these two aspects is a crucial question.

Social implications – The social impact is that copyright infringement risks of digital library is reduced; the lawsuit rate and economic loss due to copyright infringement are thereby decreased as well.

Originality/value – This paper introduces the evaluation of the potential risks of copyright infringement based on the extension theory in digital library. The results provide support for the decision-makers in handling the potential risks of copyright infringement in digital library.

Keywords Early warning, Digital library, Extension theory, Risks of copyright infringement

Paper type Research paper

The authors would like to express their sincere thanks to three anonymous referees for their constructive comments and suggestions, which were very helpful in improving the paper. This work was supported by the Ministry of Education of Humanities and Social Science project (14YJC630056) and Social Science Development Program of Fujian Province, China (2014C064).

The Electronic Library
Vol. 34 No. 2, 2016
pp. 250-264

0264-0473
DOI 10.1108/EL-04-2014-0064

Introduction

In the digital age, the digital library can reproduce, store and disseminate network information resources (NIRs) by applying new technology, which provides great convenience to users. However, Joint (2006, p. 546) pointed out that:

> […] one click of a mouse can send a file to the whole world, and the digital library users access the copyright resources in what feels like unsupervised private space for the networked nature of electronic information resources.

It means that the controlling work of digital copyright is relatively harder than that of the copyright for a printed resource. In reality, the disputes of copyright infringement have always appeared between the digital library and the copyright owners in a web environment. Furthermore, according to Duan (2012), NIRs have the following characteristics: integration of work types, complexity of copyright ownership, fuzziness of the copyright protection standards and marginalization of the copyright owners' rights. Due to these characteristics, the risks from copyright infringement of NIRs are hard to discern. Thus, the issue of risks of copyright infringement has become one of the main obstacles for digital library, and these risks may prevent digital library from creating new works and using new technologies.

To meet the above challenges of digital copyright, it is necessary to study the risks of copyright infringement. Intellectual property right (IPR) problems of digital library have been studied by several scholars from a macroscopic view (Mahesh and Mittal, 2009; Zhang, 2007). Several scholars have also studied the risks of IPR in digital library. Their researches mainly involved four aspects:

(1) IPR risks existing in digital library (Wang, 2004);

(2) factors (Xie, 2004) and types (Ran, 2009) of IPR risks;

(3) tactics to avoid IPR risks caused by the digital library or its users (Chen and Wang, 2008) through technological methods (García González and Gil, 2008; Wyatt and Hahn, 2011), management (Hugenholtz, 1994), laws (Seadle, 2008) and so forth; and

(4) methods in identifying and quantifying IPR risks.

In regard to the fourth aspect, Li and Zhang (2011) introduced the theory of risk management to research IPR risks in digital library, then exploited risk matrix analysis to evaluate risk events and provided effective strategies for digital library to control IPR risks. Zhang and Yuan (2012) assessed IPR risks based on extreme value statistics and fuzzy comprehensive appraisement. However, the data for evaluation in their research were obtained by expert scoring, which is subjective.

Recently, several scholars have begun to study the copyright risks of NIRs. For instance, Zhang *et al.* (2011) used work breakdown structure-resource barrier structure to identify the risks of copyright from the perspective of the law; Li and Zhang (2014) built an indicator system of risk assessment for copyright based on the system theory, and then, built another risk assessment indicator system based on the analysis of the factor influences from the perspective of the Internet content service provider. However, according to our research, there are few studies which were concerned with the evaluation of the risks of copyright infringement in digital library.

From the above review, it can be seen that there are few studies on the IPR problem that use objective data to study the quantitative risks of copyright infringement in digital library. Therefore, in this paper, the focus is on checking and evaluating the risks of copyright infringement in a digital library. Specifically, an early warning model used for evaluating risks is built, and an application is demonstrated to show the usefulness and effectiveness of the approach.

The rest of the paper is constructed as follows. The first section briefly reviews the methodology of extension theory. The second section constructs the indicator system and analyzes the risks of copyright infringement in a digital library. Then, an early warning model and an application with the data of the digital library of XX University are presented in the third section. A summary and conclusions are offered in the last section.

Extension theory

In this section, the extension theory and the rationale of the study are illustrated. In real-world terms, there are many problems with contradictions, goals and incompatibility conditions. It is not enough if only the quantitative relation is considered. More consideration on the matter, the characteristics and their measure together must be taken. Moreover, their relations and their changes must be considered (Cai, 1999). To solve these contradictive problems, Cai (1983) proposed the use of extension theory to sum up the expressions and treatment methods of all the contradictions and the formalization and lexicalization ways, as well as the mathematics. This theory is not only a new comprehensive evaluation method but also a qualitative and quantitative method. In line with this approach, one can use a formal model to discuss the possible developments and innovation of rules and methods. The theory has been widely applied to various fields: economy, management, artificial intelligence and so forth (Chao and Chen, 2011; Ren *et al.*, 2013; Wong and Hu, 2014).

Extension theory is a combination of matter element theory and extension set theory. In the following sub-sections, the concepts of matter element, extension set and the rationale of extension theory in solving the contradictions and incompatibility problems are introduced.

Matter element theory

In this sub-section, the matter element theory will be introduced. Cai (1983) pointed out that all of the matter in the world can be described in the following matter element model:

$$R = \{N, C, V\} = \begin{bmatrix} R_1 \\ R_2 \\ \vdots \\ R_n \end{bmatrix} = \begin{bmatrix} N_i & c_1 & v_1 \\ & c_2 & v_2 \\ & \vdots & \vdots \\ & c_n & v_n \end{bmatrix} \quad (1)$$

In equation (1), R stands for matter element, N is the name of the matter, C is the characteristic of N and V is the magnitude vector of C. N, C and V are the three fundamental elements of the model. In this paper, the ith object under evaluation is defined as $N_i(i = 1, 2, \cdots, m)$, the indicators are denoted as $c_1, c_2, \cdots, c_j, \cdots c_n$ and the value ranges of the indicators are denoted as $v_1, v_2, \cdots, v_j, \cdots v_n$.

Extension set

Definition of extension set. Let Ω be the space of objects and x be a generic element of Ω, $\widetilde{A}$ can be defined as follows:

$$\widetilde{A} = \{(x, y) | x \in \Omega, y = K(x) \in (-\infty, \infty)\} \tag{2}$$

Here, $\widetilde{A}$ is an extension set of Ω and $y = K(x)$ is the correlation function of $\widetilde{A}$, which is defined to quantify the relationship between an element and a set.

Definition of correlation function. Set $X_0 = < a, b >, X = < c, d >$ and $X_0 \in X$; then, the correlation function can be defined as follows:

$$K(x) = \begin{cases} \dfrac{-2\rho(x, X_0)}{|b - a|}, & x \in X_0 \\ \dfrac{\rho(x, X_0)}{\rho(x, X) - \rho(x, X_0)}, & x \notin X_0 \end{cases} \tag{3}$$

where,

$$\begin{aligned} \rho(x, X_0) &= \left| x - \frac{1}{2}(a + b) \right| - \frac{1}{2}(b - a) \\ \rho(x, X) &= \left| x - \frac{1}{2}(c + d) \right| - \frac{1}{2}(d - c) \end{aligned} \tag{4}$$

$\rho(x,X_0)$ means the proximity between x and X_0 and $\rho(x,X)$ means the proximity between x and X. When $K(x) > 0$, it means the degree to which x belongs to $X_0 (X_0 \in \Omega)$; when $K(x) < 0$, it describes the degree to which x does not belong to X_0; when $K(x) = 0$, it is called a zero boundary. In a special condition, when $0 < K(x) < 1$, it is called the extension domain, which means that if the conditions change, then there still is a chance for x to become a part of the set.

Research rationale

Considering the complexity, contradictions and incompatibility among factors, extension theory is introduced to evaluate the risks of copyright infringement. The basic idea is described as follows.

First, the authors establish the matter element model of copyright infringement in digital library, including the matter, the characteristics and the measure. The matter is the evaluated indicator, and the characteristics are the next level indicators.

Second, the evaluation of matter element should be determined. After selecting a digital library to study, the evaluation of matter element is the specific data of indicators corresponding to the digital library.

Next, to get early warning results of high-level indicators through the results of low-level indicators, the next step is to assign weights to the low-level indicators, which denote the importance of each indicator.

Then, the authors need to calculate the correlation function between the indicator and the degree classified by equations (3) and (4). Therefore, the degree of each indicator can be identified.

Finally, by calculating the comprehensive correlation function, the final evaluation results on the risks of copyright infringement in the digital library are calculated.

Indicator system and weights of risks of copyright infringement

Construction of indicator system

To obtain a comprehensive evaluation of the potential risks of copyright infringement in a digital library, the authors build an effective indicator system, which includes all the critical factors, such as *the threat of the risk factors*, *the vulnerability of digital library* and *the environment*. According to the literature, Li and Zhang (2014) built a comprehensive indicator system to evaluate the risks of copyright infringement for network service providers, and the indicator system is demonstrated effectively and feasibly in the practical application. They pointed out that this indicator system is a common one and can be used for all network service providers, including the digital library. Therefore, the indicator system is adopted in this paper. The low effectiveness indicators which are hard to quantify are discarded in the current research, and the remaining indicators are presented in Table I, including 3 first-level indicators, 9 second-level indicators and 22 third-level indicators.

Determination of indicators weight

Different indicators play different roles in the early warning of risks of copyright infringement in a digital library, so choosing an appropriate approach to determine the weights of the indicators is the first task. Because of the great unpredictability of the risks, the authors choose the analytic hierarchy process (AHP) (Saaty, 2013) to determine the weights. This methodology not only is suitable to deal with uncertain and subjective information but also allows decision-makers to depend on their experiences, insights and intuitions to judge problems. In this paper, the judgment matrix is attained after the experts compared the importance of one indicator to another by means of the construction of a ratio scale corresponding to the priorities of alternatives. The final weights of indicators are shown in Table I. The whole system passed the consistency examination.

Application to the digital library of XX University

In this section, an early warning degree of the risks on copyright infringement in the digital library of XX University is examined based on the extension theory. The digital library of XX University is not the most representative one but the most common one. The indicator system used in this paper is suited to digital library, especially the academic library. As it is convenient for the authors to obtain first-hand data, the digital library of XX University was chosen as the application object. At the same time, this application is being used to illustrate that this method is feasible and reasonable in evaluating the risks of copyright infringement. If the result is consistent with the reality, then this method proves to be effective in solving the evaluation of the risks.

This application is carried out by following three steps, as described below:

(1) collect the early warning data of the indicators based on the constructed indicator system;

(2) diagnose the alarm situation with the extension theory; and

(3) analyze risks of copyright infringement based on the early warning results.

First-level indicators	Weight	Second-level indicators	Weight	Third-level indicators	Weight
Vulnerability of digital library	0.309	Self risks of digital library	0.600	Service	0.185
				Risk tolerance	0.170
				Literacy of IPR	0.228
				Irregularities	0.417
		Risks of management	0.400	Perfect rate of risk policy	0.333
				Standardization of risk policy	0.222
				Implement rate of risk policy	0.444
Risk of network information resources	0.581	Risks of collecting resources	0.445	Proportion of unknown or illegal documents	0.490
				Proportion of unknown or illegal audios	0.198
				Proportion of unknown or illegal videos	0.312
		Risks of dealing with resources	0.222	Proportion of processed documents	0.539
				Proportion of processed audios	0.164
				Proportion of processed videos	0.297
		Risks of applying resources	0.111	Proportion of overused documents	0.539
				Proportion of overused audios	0.164
				Proportion of overused videos	0.297
		Risk of earning profits	0.222	Proportion of revenues from unknown or illegal documents	0.539
				Proportion of revenues from unknown or illegal audios	0.164
				Proportion of revenues from unknown or illegal videos	0.297
Risk of environment	0.110	Risks of copyright protection	0.400	Intensity of copyright protection	1
		Risks of network	0.400	Intensity of network risks	1
		Risks from users	0.200	Literacy of users	1

Table I. Early warning indicator system and weights of risks of copyright infringement

Collect the early warning data

The early warning data are collected through the Delphi method and interviews. There are eight specialists who participated in the investigation. According to the suggestions from experts, the early warning degrees are divided into five grades: *non-warning*, *light warning*, *moderate warning*, *high warning* and *serious warning*. The ranges of the value of every indicator corresponding to the five grades are also determined by the experts. The classification is demonstrated in Table II.

The values of the digital library of XX University are also displayed in Table II. The major data are gathered from the chief librarian through an interview, and the other data are obtained from the literature and the report. For the indicator of *copyright protection intensity*, the value of 4.45 is gained through a predictive calculation (Peng, 2012).

Table II.
The classification of early warning degrees and values of the digital library of XX university

Projects	Indicators of CIR[a]	Quantification of indicators	Classification of early warning degrees					Value
			Non-warning	Light warning	Moderate warning	High warning	Serious warning	
Vulnerability of digital library	Service	Scores of service[b]	$2 \le M < 5$	$5 \le M < 6$	$6 \le M < 7$	$7 \le M < 8$	$8 \le M \le 10$	2.6
	Risk tolerance	The maximal loss acceptable/the total assets	$0.4 \le M < 1$	$0.3 \le M < 0.4$	$0.2 \le M < 0.3$	$0.1 \le M < 0.2$	$0 < M < 0.1$	0.12
	Literacy of IPR	Number of staffs with IPR consciousness/the total number of staffs	$0.8 \le M \le 1$	$0.6 \le M < 0.8$	$0.4 \le M < 0.6$	$0.2 \le M < 0.4$	$0 \le M < 0.2$	0.23
	Irregularities	Number of staffs with irregularities/the total number of staffs	$0 \le M < 0.15$	$0.15 \le M < 0.3$	$0.3 \le M < 0.45$	$0.45 \le M < 0.6$	$0.6 \le M \le 1$	0.42
	Perfect rate of risk policy	Number of practical CIR policies/ the total number of CIR policies	$0.8 \le M < 1$	$0.7 \le M < 0.8$	$0.6 \le M < 0.7$	$0.4 \le M < 0.6$	$0 < M < 0.4$	0.9
	Standardization of risk policy	Number of standard CIR policies/ the total number of CIR policies	$0.8 \le M < 1$	$0.7 \le M < 0.8$	$0.6 \le M < 0.7$	$0.4 \le M < 0.6$	$0 < M < 0.4$	0.85
	Implement rate of risk policy	Number of implemented CIR policies/the total number of CIR policies	$0.95 \le M < 1$	$0.9 \le M < 0.95$	$0.8 \le M < 0.9$	$0.65 \le M < 0.8$	$0 < M < 0.65$	0.96
Risks of network information resources	Proportion of unknown or illegal documents	Number of unknown or illegal documents/the total number of documents	$0 \le M < 0.1$	$0.1 \le M < 0.15$	$0.15 \le M < 0.25$	$0.25 \le M < 0.4$	$0.4 \le M < 1$	0.05
	Proportion of unknown or illegal audios	Number of unknown or illegal audios/the total number of audios	$0 \le M < 0.15$	$0.15 \le M < 0.25$	$0.25 \le M < 0.35$	$0.35 \le M < 0.5$	$0.5 \le M < 1$	0.1
	Proportion of unknown or illegal videos	Number of unknown or illegal videos/the total number of videos	$0 \le M < 0.15$	$0.15 \le M < 0.25$	$0.25 \le M < 0.35$	$0.35 \le M < 0.45$	$0.45 \le M < 1$	0.2
	Proportion of processed documents	Number of processed documents/ the total number of documents	$0 \le M < 0.1$	$0.1 \le M < 0.2$	$0.2 \le M < 0.4$	$0.4 \le M < 0.6$	$0.6 \le M < 1$	0.13
	Proportion of processed audios	Number of processed audios/the total number of audios	$0 \le M < 0.15$	$0.15 \le M < 0.3$	$0.3 \le M < 0.5$	$0.5 \le M < 0.7$	$0.7 \le M < 1$	0.45
	Proportion of processed videos	Number of processed videos/the total number of videos	$0 \le M < 0.15$	$0.15 \le M < 0.3$	$0.3 \le M < 0.5$	$0.5 \le M < 0.65$	$0.65 \le M < 1$	0.32
	Proportion of overused documents	Number of overused documents/the total number of documents	$0 \le M < 0.05$	$0.05 \le M < 0.1$	$0.1 \le M < 0.15$	$0.15 \le M < 0.25$	$0.25 \le M < 1$	0.12
	Proportion of overused audios	Number of overused audios/the total number of audios	$0 \le M < 0.1$	$0.1 \le M < 0.15$	$0.15 \le M < 0.2$	$0.2 \le M < 0.35$	$0.35 \le M < 1$	0.05
	Proportion of overused videos	Number of overused videos/the total number of videos	$0 \le M < 0.1$	$0.1 \le M < 0.15$	$0.15 \le M < 0.2$	$0.2 \le M < 0.3$	$0.3 \le M < 1$	0.12
	Proportion of revenues from unknown or illegal documents	Revenues from unknown or illegal documents/the total revenues from documents	$0 \le M < 0.1$	$0.1 \le M < 0.2$	$0.2 \le M < 0.35$	$0.35 \le M < 0.55$	$0.55 \le M < 1$	0.05

(*continued*)

Table II.

Projects	Indicators of CIR[a]	Quantification of indicators	Classification of early warning degrees: Non-warning	Light warning	Moderate warning	High warning	Serious warning	Value
	Proportion of revenues from unknown or illegal audios	Revenues from unknown or illegal audios/the total revenues from audios	$0 \le M < 0.15$	$0.15 \le M < 0.3$	$0.3 \le M < 0.45$	$0.45 \le M < 0.65$	$0.65 \le M < 1$	0.1
	Proportion of revenues from unknown or illegal videos	Revenues from unknown or illegal videos/the total revenues from videos	$0 \le M < 0.15$	$0.15 \le M < 0.3$	$0.3 \le M < 0.45$	$0.45 \le M < 0.6$	$0.6 \le M < 1$	0.2
Risks of environment	Intensity of copyright protection	Legislation strength × enforcement strength	$5.5 < M \le 6$	$4.5 < M \le 5.5$	$3.5 < M \le 4.5$	$1.6 < M \le 3.5$	$0 < M \le 1.6$	4.45
	Intensity of network risks	Times of site attacked successfully/the total times attacked	$0 \le M < 0.005$	$0.005 \le M < 0.015$	$0.015 \le M < 0.025$	$0.025 \le M < 0.04$	$0.04 \le M < 1$	0.001
	Literacy of users	Number of netizens with university degree or above/the total number of netizens	$0.45 \le M < 1$	$0.35 \le M < 0.45$	$0.2 \le M < 0.35$	$0.1 \le M < 0.2$	$0 < M < 0.1$	0.21[c]

Notes: [a] CIR means copyright infringement risks; [b] service is abbreviated as S: service of network content (S1) with ten points, provides information storage space (S2) with eight points, service of software and technology (S3) with five points, provides network resources accessed copyright (S4) with two points, then $S = \Sigma w_i S_i$; [c] the 31th Internet report: Internet users and structural characteristics. [EB/OL]. (2013-01-15) [2014-01-06] http://it.people.com.cn/n/2013/0115/c354305-20207435-2.html

Diagnose the alarm situation

In this section, the alarm situation of the digital library of XX University will be diagnosed. The indicator of *self risks* is discussed in detail, which is taken as an example to illustrate how to apply extension theory to forecast the risks of copyright infringement.

The early warning of the third-level indicators

Establish the matter element model. The matter elements R_{0i} of the classical domain of the digital library of XX University are defined as follows:

$$R_{01} = \begin{bmatrix} \text{Non} - \text{warning} & \text{Service} & \langle 2,\ 5\rangle \\ & \text{Risk tolerance} & \langle 0.4,\ 1\rangle \\ & \text{Literacy of IPR} & \langle 0.8,\ 1\rangle \\ & \text{Irregularities} & \langle 0,\ 0.15\rangle \end{bmatrix},$$

$$R_{02} = \begin{bmatrix} \text{Light warning} & \text{Service} & \langle 5,\ 6\rangle \\ & \text{Risk tolerance} & \langle 0.3,\ 0.4\rangle \\ & \text{Literacy of IPR} & \langle 0.6,\ 0.8\rangle \\ & \text{Irregularities} & \langle 0.15,\ 0.3\rangle \end{bmatrix}$$

$$R_{03} = \begin{bmatrix} \text{Moderate warning} & \text{Service} & \langle 6, 7\rangle \\ & \text{Risk tolerance} & \langle 0.2, 0.3\rangle \\ & \text{Literacy of IPR} & \langle 0.4, 0.6\rangle \\ & \text{Irregularities} & \langle 0.3, 0.45\rangle \end{bmatrix},$$

$$R_{04} = \begin{bmatrix} \text{High warning} & \text{Service} & \langle 7, 8\rangle \\ & \text{Risk tolerance} & \langle 0.1, 0.2\rangle \\ & \text{Literacy of IPR} & \langle 0.2, 0.4\rangle \\ & \text{Irregularities} & \langle 0.45, 0.6\rangle \end{bmatrix},$$

$$R_{05} = \begin{bmatrix} \text{Serious warning} & \text{Service} & \langle 8, 10\rangle \\ & \text{Risk tolerance} & \langle 0, 0.1\rangle \\ & \text{Literacy of IPR} & \langle 0, 0.2\rangle \\ & \text{Irregularities} & \langle 0.6, 1\rangle \end{bmatrix}$$

And the matter element of segment domain R_p is:

$$R_p = \begin{bmatrix} N_p & \text{Service} & \langle 2, 10\rangle \\ & \text{Risk tolerance} & \langle 0, 1\rangle \\ & \text{Literacy of IPR} & \langle 0, 1\rangle \\ & \text{Irregularities} & \langle 0, 1\rangle \end{bmatrix}$$

Determine the matter element for early warning. Based on the investigation, the matter element R for the early warning of the digital library of XX University is defined as follows:

$$R = \begin{bmatrix} N & \text{Service} & 2.6 \\ & \text{Risk tolerance} & 0.12 \\ & \text{Literacy of IPR} & 0.23 \\ & \text{Irregularities} & 0.42 \end{bmatrix}$$

Assign the weights to the indicators of the risks of copyright infringement. The weights of the third-level indicators that belong to *self risks* are gained by way of AHP:

$$W_{111} = (0.185, 0.170, 0.228, 0.417)$$

Calculate the correlation function. According to equations (3) and (4), the correlation function between the indicator and degree can be obtained. Take the *service* in the indicators for example, the authors will elaborate and illustrate the way to calculate the correlation function.

① The value of the *service* is 2.6, that is, $v_{1111} = 2.6 \in [2, 5]$, which means the level of the infringement risk of *service* belongs to non-warning and the proximity between 2.6 and [2, 5] is:

$$\rho(2.6, \langle 2, 5\rangle) = \left| 2.6 - \frac{1}{2}(2 + 5) \right| - \frac{1}{2}(5 - 2) = -0.600,$$

Then, the correlation function of 2.6 and the level of non-warning is:

$$K_{1111}(2.6) = \frac{-2\rho(2.6, \langle 2, 5\rangle)}{5 - 2} = 0.400$$

② $v_{1111} = 2.6 \notin [5, 6]$, and the proximity between 2.6 and [5, 6] is:

$$\rho(2.6, \langle 5, 6\rangle) = \left| 2.6 - \frac{1}{2}(5 + 6) \right| - \frac{1}{2}(6 - 5) = 2.400,$$

The proximity between 2.6 and [2, 10] is:

$$\rho(2.6, \langle 2, 10\rangle) = \left| 2.6 - \frac{1}{2}(2 + 10) \right| - \frac{1}{2}(10 - 2) = -0.600,$$

Then, the correlation function of 2.6 and the level of light warning is:

$$K_{1112}(2.6) = \frac{\rho(2.6, \langle 2, 5\rangle)}{\rho(2.6, \langle 2, 10\rangle) - \rho(2.6, \langle 2, 5\rangle)} = -0.800.$$

③ As $v_{1111} = 2.6 \notin [6, 7]$, $v_{1111} = 2.6 \notin [7, 8]$ and $v_{1111} = 2.6 \notin [8, 10]$, following the same steps in ②, we have:

$$K_{1113}(2.6) = -0.850,\ K_{1114}(2.6) = -0.880,\ K_{1115}(2.6) = -0.900.$$

The correlation functions between 2.6 and the five grades compose the first component of K_{111} denoted as $K_{111}^{(1)}$, that is:

$$K_{111}^{(1)} = (0.400, -0.800, -0.850, -0.880, -0.900),$$

The corresponding level with a positive number is non-warning, so the level of *service* belongs to non-warning. Following the above computations, the correlation functions of the other three indicators in each degree are provided, so the incidence matrix is:

$$K_{111} = \begin{bmatrix} 0.400 & -0.800 & -0.850 & -0.880 & -0.900 \\ -0.700 & -0.600 & -0.400 & 0.400 & -0.143 \\ -0.713 & -0.617 & -0.425 & 0.300 & -0.115 \\ -0.391 & -0.222 & 0.400 & -0.067 & -0.300 \end{bmatrix}.$$

Calculate the comprehensive correlation function. Based on the incidence matrix K_{111} and the weights of *service*, *risks tolerance*, *literacy of IPR* and *irregularities*, the comprehensive correlation function of *self risks* can be calculated as follows:

$$K_{11} = W_{111} * K_{111} = (-0.371, -0.484, -0.156, -0.054, -0.342).$$

In the same way, we can obtain the correlation functions of risks of management, *collecting resources*, *dealing with resources*, *applying resources*, *earning profits*, *copyright protection*, *network* and *users*. The results are presented as follows:

$K_{12} = (0.622, -0.311, -0.600, -0.744, -0.838)$, $K_{21} = (0.560, 0.001, -0.508, -0.667, -0.761)$,

$K_{22} = (-0.270, 0.265, -0.047, -0.487, -0.632)$, $K_{23} = (0.553, -0.060, -0.546, -0.673, -0.793)$,

$K_{24} = (0.589, -0.126, -0.613, -0.755, -0.827)$, $K_{31} = (-0.404, -0.031, 0.100, -0.380, -0.648)$,

$K_{32} = (0.400, -0.800, -0.933, -0.960, -0.975)$, $K_{33} = (-0.533, -0.400, 0.133, -0.045, -0.344)$.

From the above results, we know that *the risks of management*, *collecting resources*, *applying resources*, *earning profits* and *network* all belong to the level of non-warning with little risks, and *risks of dealing with resources* belongs to light warning. However, *copyright protection risks* and *risks from users* belong to moderate warning, which are uncontrollable risks for the digital library, and there are certain risks involving copyright disputes.

The early warning of the second-level indicators

It is easy to find out the results of the third-level indicators composed on the early warning matrix of the second level. Thus, the correlation functions of *vulnerability of digital library*, *NIRs risks* and *environment risks* can be calculated as follows:

$$\begin{aligned} K_1 &= (0.600, 0.400)\begin{bmatrix} -0.371 & -0.484 & -0.156 & -0.054 & -0.342 \\ 0.622 & -0.311 & -0.600 & -0.744 & -0.838 \end{bmatrix}, \\ &= (0.027, -0.415, -0.333, -0.330, -0.541) \end{aligned}$$

$$K_2 = (0.444, 0.222, 0.111, 0.222)\begin{bmatrix} 0.560 & 0.001 & -0.508 & -0.667 & -0.761 \\ -0.270 & 0.265 & -0.047 & -0.487 & -0.632 \\ 0.553 & -0.060 & -0.546 & -0.673 & -0.793 \\ 0.589 & -0.126 & -0.613 & -0.755 & -0.827 \end{bmatrix},$$

$$= (0.381, 0.025, -0.433, -0.647, -0.750)$$

$$K_3 = (0.400, 0.400, 0.200)\begin{bmatrix} -0.404 & -0.031 & 0.100 & -0.380 & -0.648 \\ 0.400 & -0.800 & -0.933 & -0.960 & -0.975 \\ -0.533 & -0.400 & 0.133 & -0.045 & -0.344 \end{bmatrix}$$

$$= (-0.108, -0.413, -0.307, -0.545, -0.718)$$

The above results show that the indicators of the second level all belong to the non-warning degree. It means that the three second-level indicators of the risks of copyright infringement in the digital library of XX University are less likely to be infringed.

The early warning of the first-level indicators

The calculation of the early warning process for the first level is similar to the second one, and then, the result of the risks of copyright infringement in the digital library of XX University is:

$$K_P = (0.309, 0.581, 0.11)\begin{bmatrix} 0.027 & -0.415 & -0.333 & -0.330 & -0.541 \\ 0.381 & 0.025 & -0.433 & -0.647 & -0.750 \\ -0.108 & -0.413 & -0.307 & -0.545 & -0.718 \end{bmatrix},$$

$$= (0.218, -0.159, -0.388, -0.538, -0.682)$$

Thus, the early warning degree of the risks of copyright infringement in the digital library of XX University is non-warning.

Discussion

Using the early warning model constructed above, the authors quantified the risks of copyright infringement in the digital library of XX University. From the above calculation, it is determined that the digital library of XX University has a slim chance of infringing copyright. This result is consistent with the reality known from the chief librarian by the way of questionnaires and interviews.

There are several reasons for the result. First, the digital library of XX University is a non-profit unit, which means its risks of copyright infringement are much lower than the for-profit ones. Second, 80 per cent of the network resources in this digital library are authorized copyright through purchase or in other ways – for example, signing an agreement with the copyright owners – all of which may greatly reduce the possibility of infringement. Conversely, only 10 per cent of the network resources need to be collected or dealt with before providing users access to them. Third, the other 10 per cent of resources and services are software tools, searching and linking services. The software tools are the most commonly used ones and are free tools available on the Internet; the services of searching and linking are applicable to safe harbour principles. Fourth, the digital library provides services for the teachers and students at XX University, which significantly lowers the scope of infringement.

In a word, there are few risks of copyright infringement for both the service mode and service contents in the digital library of XX University. Therefore, the model proposed is effective in measuring the risks of the copyright infringement in a digital library to some extent.

Conclusion

To measure the risks of the copyright infringement in the digital library, a method is put forward. More specifically, the extension theory is introduced in the study to construct an early warning model. This model gives the digital library intuitive insight into the existing risks of copyright infringement.

According to the study, it is known that:

- the early warning extension theory model is effective in distinguishing the degree of the potential risks of the copyright infringement in a digital library; and
- the ranges of the value and the values of the indicators can directly affect the results while using this approach, so the accuracy of these two aspects is a crucial question.

The contribution of this study can be summarized into three aspects:

(1) As a practical influence, the research can be used to help a digital library to predict potential copyright risks in the process of daily work, so that economic loss can be reduced to a certain degree.

(2) In a research aspect, the research can broaden the application scope of the extension theory, and it uses data to analyze copyright infringement risks of a digital library, which is not just confined to theory analysis.

(3) The social impact is that copyright infringement risks of a digital library are reduced; the lawsuit rate and economic loss due to copyright infringement are thereby decreased as well.

References

Cai, W. (1983), "Extension set and non-compatible problems", *Science Exploration*, Vol. 1 No. 1, pp. 25-29.

Cai, W. (1999), "Extension theory and its application", *Chinese Science Bulletin*, Vol. 44 No. 17, pp. 1538-1548.

Chao, K.H. and Chen, J.W. (2011), "State-of-health estimator based-on extension theory with a learning mechanism for lead-acid batteries", *Expert Systems with Applications*, Vol. 38 No. 12, pp. 15183-15193.

Chen, C.F. and Wang, Y.D. (2008), "Risk management strategy of intellectual property for library users", *Library Tribune*, Vol. 28 No. 6, pp. 16-20.

Duan, W. (2012), "Outline of network copyright protection", *Outline of Publishing Culture and Industry Series of Online Copyright Protection of Cultural Resources and Industry Library*, Central China Normal University Press, Wuhan, pp. 1-3.

García González, R. and Gil, R. (2008), "Semantic copyright management for Internet-wide knowledge sharing and reuse", *Online Information Review*, Vol. 32 No. 5, pp. 585-595.

Hugenholtz, P.B. (1994), "Copyright and electronic document delivery services", *Interlending & Document Supply*, Vol. 22 No. 3, pp. 8-14.

Joint, N. (2006), "Risk assessment and copyright in digital libraries", *Libraries Review*, Vol. 55 No. 9, pp. 545-548.

Li, C. and Zhang, W.D. (2011), "Research on digital library intellectual property risk management", *Information Studies: Theory & Application*, Vol. 34 No. 11, pp. 31-35.

Li, C. and Zhang, W.D. (2014), "A new risk assessment indicator system of copyright for network information resources", *Documentation, Information & Knowledge*, Vol. 27 No. 2, pp. 102-110.

Mahesh, G. and Mittal, R. (2009), "Digital content creation and copyright issues", *The Electronic Library*, Vol. 27 No. 4, pp. 676-683.

Peng, H. (2012), "Index system building and intensity calculation of copyright protection", *Theory and Empirical Research of Copyright Protection System*, Academy of Social Sciences Press, Shanghai, pp. 16-36.

Ran, C.J. (2009), "Research on the main risk types of intellectual property in library", *Library and Information Service*, Vol. 53 No. 11, pp. 28-32.

Ren, J., Manzardo, A., Toniolo, S. and Scipioni, A. (2013), "Sustainability of hydrogen supply chain Part II: prioritizing and classifying the sustainability of hydrogen supply chains based on the combination of extension theory and AHP", *International Journal of Hydrogen Energy*, Vol. 38 No. 32, pp. 13845-13855.

Saaty, T.L. (2013), "The modern science of multicriteria decision making and its practical applications: the AHP/ANP approach", *Operations Research*, Vol. 61 No. 5, pp. 1101-1118.

Seadle, M. (2008), "Copyright in the networked world: gray copyright", *Library Hi Tech*, Vol. 26 No. 2, pp. 325-332.

Wang, H. (2004), "Risks of intellectual property in digital library and risk aversion", *Theory and Practice of Library*, Vol. 19 No. 6, pp. 12-13.

Wong, H. and Hu, B.Q. (2014), "Application of improved extension evaluation method to water quality evaluation", *Journal of Hydrology*, Vol. 509 No. 2, pp. 539-548.

Wyatt, A.M. and Hahn, S.E. (2011), "Copyright concerns triggered by Web 2.0 uses", *Reference Services Review*, Vol. 39 No. 2, pp. 303-317.

Xie, Y. (2004), "Analysis of risk factors of intellectual property in digital library", *Research in Library Science*, Vol. 26 No. 10, pp. 98-100.

Zhang, K., Li, Y. and Guo, W. (2011), "Research on the influence of intellectual property risk on cooperative behavior in R&D", *Management Review*, Vol. 23 No. 12, pp. 76-83.

Zhang, W. (2007), "Digital library intellectual property right evaluation and method", *The Electronic Library*, Vol. 25 No. 3, pp. 267-273.

Zhang, W.D. and Yuan, Y. (2012), "The evaluation of intellectual property risk in digital library", *Information Studies: Theory & Application*, Vol. 35 No. 5, pp. 28-32.

Further reading

Wang, M.H., Chung, Y.K. and Sung, W.T. (2009), "The fault diagnosis of analog circuits based on extension theory[M]", *Emerging Intelligent Computing Technology and Applications. Springer, Heidelberg*, pp. 735-744.

About the authors

Chan Li gained a Master's degree and a PhD degree in Information Science from Fuzhou University. She is a Lecturer in the Department of the College of Public Administration of Fujian Agriculture and Forestry University. Her current research interests include information

management and intellectual property, as well as risk management. Li has published research papers in the areas of digital libraries, intellectual property, knowledge and management in Chinese periodicals and scholarly journals. Chan Li is the corresponding author and can be contacted at: shandongweifanglch@163.com

Wen-De Zhang received his BE in Mechanical Manufacturing Engineering Equipments and Automation from Fuzhou University with advanced studies in Information Science from Wuhan University. He has an MA in Business Administration from Hong Kong Open University and a PhD in Management Science and Engineering from Huazhong University of Science and Technology. Currently, he is working at Fuzhou University as a Professor. His research interests are mainly in the areas of information management, intellectual property, information processing and computer applications. He has written or edited 12 books and 160 research papers, which were published in periodicals and scholarly journals in China. In addition, he has obtained four patents relevant to library utilities. Professor Zhang is as well known for his dedication to teaching and mentoring students as he is for his research. He is a Mentor of Master's and doctoral students in Information Science.

Yi-Xin Lan is a Lecturer in the School of Economics and Management of Fuzhou University. He holds a PhD of Management Science and Engineering and a MSC in College of Mathematics and Computer Science of Fuzhou University. His current research interests are mathematical statistics and data envelopment analysis. He has published many refereed articles in international journals in the areas of mathematics, economics and management, in such titles as *Mathematical and Computer Modelling*, *Economic Modelling*, *Expert Systems with Applications*, *Communications in Statistics-Theory and Methods* and so on.