한국산학기술학회논문지
Vol. 9, No. 5, pp. 1255-1259, 2008

사례중심의 프로그램 유사성 및 도용여부에 관한연구

김용원[1], 이성훈[2*]

A Study for Program Similarity on Case Study

Yoeng-Won Kim[1] and Seong-Hoon Lee[2*]

요 약 본 연구는 실제 사례를 중심으로 하여 프로그램의 복제정도 및 도용여부를 판단하기 위한 모형을 개발하는 내용이다. 본 논문에서는 복제정도 및 도용여부를 판별하기 위한 항목을 기술하였다. 또한 각 항목들은 각각의 중요도가 다르기 때문에 각 항목별 중요도를 결정하고 각 항목에 대해 복제정도를 산출하였다. 이를 기반으로 하여 프로그램간의 전체적인 복제여부를 판단한 연구내용이다.

Abstract This paper deals with a study on copy rate for program when there is a conflict with programs. In this paper, we presented a setting method on weight and degree of importance on each factors because degree of importance on one factor was different from degree of importance onother factors. Based on various factors, we judged on copy rate for program.

1. 서 론

정부의 각종지원 프로그램에 의해 소프트웨어 산업은 양적, 질적으로 많은 성장과 변화를 가져왔다. 이러한 고도성장에 따른 문제점중 하나는 소프트웨어에 대한 지적재산권 문제이며 이로 인해 많은 분쟁이 일어나고 있다[2[5]]. 따라서 지식정보사회에서 지적재산권에 대한 공정한 분쟁의 해결을 위하여 제도개선 및 체계화가 중요한 문제로 대두되고 있다[5]. 본 연구에서는 GIS 프로그램에 실제 사례를 중심으로 하여 프로그램의 유사성 및 도용성 정도를 도출할 수 있는 방안을 제시하고자 한다. 프로그램에 대한 유사성이란 프로그램간에 소스코드 및 데이터 구조등이 얼마나 유사한지를 나타내는 것이라 할 수 있다. 따라서 프로그램에 사용된 자료구조, 프로그램에 대한 소스코드, 사용자 인터페이스등이 유사정도를 측정하기 위한 척도로 사용되어 왔다. 본 연구에서는 이러한 평가척도를 개발하고 각 평가척도에 대한 주요성을 반영시키기 위한 가중치를 설정하였다. 2장에서는 사례분석에 대한 개요 및 감정기준에 대해 기술한다. 3장에서는 감정수행방법을 마지막으로 4장에서 결론을 기술하였다.

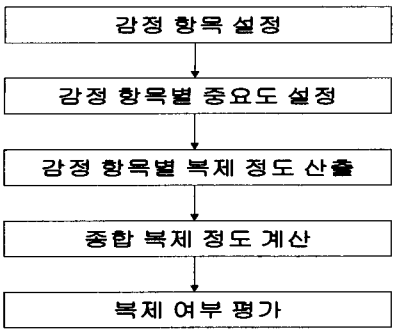
2. 연구개요

본 연구의 대상은 실제 사례로서 다음과 같은 상황에서 발생되었다. 관련회사에 대해서는 가명으로 하여 기술한다. A라는 회사는 AA라는 GIS 프로그램을 개발하여 시판중인 상황에서 B라는 회사가 BB라는 GIS 프로그램을 시판하기 시작하였다. 여기서 A라는 회사는 B라는 회사가 개발하였다고 하는 BB 프로그램이 자사의 AA라는 프로그램을 복제하였다고 하여 중앙법원에 이의를 제기하였다. A라는 회사의 이의제기 주 내용은 다음과 같다.

- ▸ 자사의 AA 프로그램과 B사의 BB 프로그램에 대해 다음 4가지 사항의 유사성 및 도용여부 의뢰
  - Import / Export
  - 공간 연산
  - 저장관리자
  - Display

따라서 본 연구에서는 4가지 이의제기 사항들에 대해 양사의 프로그램을 다음과 같은 감정기준을 정하여 분석하였으며 [그림 1]과 같은 절차로 진행하였다.

[1]건양대학교 기업정보관리학과
[2]백석대학교 정보통신학부
*교신저자: 이성훈 (shlee@bu.ac.kr)
접수일 08년 07월 28일 수정일 08년 09월 04일 게재확정일 08년 10월 16일



[그림 1] 연구수행 절차

- 본 프로그램 감정은 고소인(A사)이 피고소인(B사)을 상대로 프로그램의 유사성 및 도용 여부에 대해 감정을 의뢰하였으므로 피고소인측에서 얼마나 고소인측의 프로그램을 인용 및 도용하였는지를 조사, 분석하며 분석대상은 위에서 기술된 4가지 사항에 국한 함.

  - Import/Export 부문 - 공간 연산 부문
  - 저장 관리 부문 - Display 부문

- 감정항목의 내용이 공개 소프트웨어를 이용한 점이 인정되면 복제정도를 "0"로 함

3. 감정수행 내용

본 연구에서의 내용은 위에서 기술된 [그림 1]의 순으로 진행되었으며 이를 단계별로 구분하여 기술한다.

3.1 감정항목 결정

먼저 제출된 두 개의 프로그램 유사성 및 도용여부 분석을 위해 감정해야 할 항목을 결정하는 단계로 다음과 같이 감정 항목을 선정하여 감정을 수행하였다.

1) 소스 프로그램(source program) : 감정대상인 4개 부문에 대한 소스코드를 비교, 분석함에 있어 아래 사항을 대상으로 비교, 분석하였음
   - 코드의 일치성
   - 프로그램 구성 방법

2) 저장관리 형태
   - 인덱스 구조 : 데이터 접근을 위한 인덱스 구조
   - 자료구조 : 파일내 필드 동일성 분석

3) 공간 연산 형태 : 대부분 공간 연산을 위한 프로그램들이 공개된 내용들이 대부분이므로 이들 연산들이 두 회사의 프로그램에 어떻게 이용되었는지를 비교, 분석
   - 공간연산 공개 소프트웨어 사용 형태

4) 화면 구성 : 화면에 디스플레이되는 아래 사항들에 대해 비교, 분석
   - 구성 메뉴
   - 구성된 아이콘 형태

5) 기타
   - 파일명의 동일성

3.2 감정항목별 중요도 설정

전 단계에서 기술한 감정 항목들에 대해 상대적인 중요도를 부여한 단계로서 각각의 감정항목들에 대해 다음과 같은 원칙하에 중요도를 부여하였다.

1) 소스 프로그램(source program) : 프로그램의 유사성 및 도용여부를 판단하기 위해서는 먼저 소스 코드 수준에서의 비교, 분석이 가장 확실하다는 측면에서 중요하다. 소스 코드의 내용이 동일하다면 이는 명확하게 도용되었다라고 단정할 수 있기 때문이다. 소스코드 수준에서 도용여부가 결정된다면 다른 감정항목에 대한 비교, 분석이 필요치 않을 것이다. 따라서 전체 중요도중에서 50%의 가중치를 부여하였다. 소스코드 항목을 세분화하면 다음과 같다.

   - 코드 일치정도 : 감정 대상물에 대한 소스코드의 일치 정도
     · 복제 및 도용여부 판단의 가장 확실한 요소
     · 중요도 : 40% 부여
   - 프로그램 구성 방법 : 프로그래밍 기법의 유사성 정도
     · 감정기법등이 알려지면서 프로그램을 간단히 수정하거나 일부 교정하는 경우 발생
     · 동일 내용을 서로 다른 사람이 프로그래밍하는 경우 프로그래밍 방법은 상이한 점이 일반적인 경향
     · 따라서 프로그래밍 방법이 동일하다면 유사성이 있는 것으로 판단
     · 중요도 : 10% 부여

2) 저장관리 형태 : 본 감정물은 GIS와 관련된 것으로 많은 공간 및 비공간 데이터들을 저장, 관리한다는 측면에서 이들에 대한 관리 방법들을 비교, 분석하는 것은 중요하다. 따라서 이에 대한 중요도는 15%를 부여하였으며 세부 내용은 다음과 같다.
- 인덱스 구조
  · 공간 및 비공간 데이터 접근을 위한 인덱스 구조 비교
  · 중요도 : 5% 부여
- 자료구조 동일성(10)
  · 레코드 필드들의 유사성 비교
  · 중요도 : 10% 부여

3) 공간 연산 : 공간 데이터들에 대한 연산(합집합(Union), 교집합(Intersection), 차집합(Difference), 선분객체 합병(Join))과 관련된 내용을 비교, 분석하는 부문으로 이들 연산들은 대부분 공개 소프트웨어에 해당한다. 따라서 본 감정물에서는 공개 소프트웨어를 사용한 형태를 비교하며 중요도는 15%로 부여하였다.

4) 화면 구성 : 사용자들을 위해 화면에 디스플레이되는 내용을 비교, 분석하며 이부문의 중요도는 15%를 부여하였다. 화면구성에 대한 세부 내용은 다음과 같다.

- 메뉴 구성
  · 프로그램에서 제공되는 기능들에 대한 메뉴 구성을 비교
  · 메뉴 구성의 동일성 정도에 따라 프로그램의 유사성을 판별
  · 중요도 : 10% 부여
- 아이콘 구성 및 형태
  · 아이콘의 구성 및 형태에 대한 동일성 정도 역시 프로그램 유사성을 판별할 수 있는 척도
  · 중요도 : 5% 부여

5) 기타 : 본 감정물의 유사성 및 도용 여부 판별을 위해 기타 사항들에 대해 조사, 분석
- 조사, 분석 내용
  · 프로그램 화일명의 동일성
- 중요도 : 5% 부여

위에서 기술된 감정 항목 및 항목별 중요도를 종합, 기술하면 다음 [표 1]과 같다.

[표 1] 감정항목 및 항목별 중요도

감정 항목	세부 내용	중요도 (%)	중요도 합(%)
소스 프로그램	프로그램 구성 방법	10	50
	코드 일치정도	40	
저장관리 형태	인덱스 구조	5	15
	자료구조 동일성	10	
공간연산	공개소프트웨어 이용형태	15	15
화면 구성	구성 메뉴	10	15
	아이콘 구성 형태	5	
기타	파일의 동일성	5	5
합 계		100	100

3.3 감정항목별 복제정도 산출

위에서 결정된 감정 항목별 복제 정도는 다음과 같은 원칙으로 계산한다. 단, 감정항목이 공개 소프트웨어를 이용한 경우 이전에 기술된 감정기준에 준하여 해당항목에 대한 복제정도를 "0"로 하였다.

• 소스코드에 대한 복제 정도
- 두 개의 프로그램간 라인 단위로 동일 내용을 추출하여 계산
- 예) A사 프로그램 라인 수 : 100, B사의 프로그램 중 동일한 라인 수 : 10
  ⇒ 복제 정도 = 10%

• 저장관리 형태에 대한 복제 정도
- GIS 관련 공간 및 비공간 데이터 접근을 위한 인덱스 구조
  예) A사 인덱스 구조 : 공간=R-트리, 비공간=B-트리
  B사 인덱스 구조 : 공간=R-트리, 비공간=B-트리
  ⇒ 복제 정도 : 100%
  ⇒ 공개 소프트웨어 이용시 복제 정도 : 0%
- 파일인 경우 : 레코드 구조의 동일성으로 계산
  예) 필드의 개수 : 20개
  동일한 필드 개수 : 5개
  ⇒ 복제 정도 : 25%

• 공간 연산에 대한 복제 정도
- 공간연산에 대한 이용형태의 동일성으로 계산
- 예) A사 : 4개 연산, B사 : 3개 연산,
  이용형태 동일 개수 : 2개 연산
  ⇒ 복제 정도 : 50%
  ⇒ 공개 소프트웨어 이용시 복제 정도 : 0%

- 화면 구성에 대한 복제 정도
  - 주 메뉴 및 부속 메뉴의 동일성으로 계산
    예) 메뉴 항목 : 10개, 유사메뉴 개수 : 2개
    ⇒ 복제 정도 : 20%
  - 사용된 아이콘의 동일성 정도로 계산
    예) 아이콘 개수 : 10개, 유사아이콘 개수 : 2개
    ⇒ 복제 정도 : 20%

- 기타
  - 파일명의 동일성
  - 감정의뢰 부분(4개 부분)에 대해 고소인측(A사)에서 제공한 576개 프로그램 파일들중에서 피고소인(B사)측 프로그램 파일들의 파일명 동일성 조사
    예) 파일 개수 : 20개, 동일 화일명 개수 : 10개
    ⇒ 복제 정도 : 50%

위에서 기술한 기준에 준하여 각 감정항목별 복제정도를 산출하면 다음 [표 2]와 같다.

[표 2] 감정항목별 복제 정도

감정항목	세부항목	원본	사본	일치	복제정도	비고
소스프로그램	프로그램 구성방식			완전상이	0%	
	코드 일치정도	48,299	44,710		0%	
저장관리형태	인덱스 구조	R-트리, B-트리	R-트리, B-트리	동일	0%	공개소스
	자료구조					
공간연산	연산이용 형태	기능 포함	외부공개 소스이용	상이	0%	공개소스
화면구성	구성메뉴	119	109	39	33%	
	아이콘 구성	40	42	13	33%	
기타	파일,함수 명명방식	576	64	31	5%	
합계		100				

3.4 종합 복제정도 계산

위에서 기술된 3.2절 및 3.3절의 내용을 기반으로 하여 두 개의 감정물에 대한 종합 복제 정도를 계산하는 단계로 본 감정 프로그램들에 대한 종합적인 복제 정도는 다음과 같은 원리에 의해 산출하였다.

- 감정 항목들에 대해 사본에 원본의 내용이 복제된 정도를 합산
- 계산시 감정 항목들의 중요도를 반영하여 계산

$$\text{종합복제정도} = \sum_{i=1}^{n} (\text{항목별중요도} \times \text{항목별복제정도}) \quad ..(1)$$

(n : 감정 항목 개수)

궁극적으로 복제정도는 위의 수식 (1)에 의해 도출된 값이 복제의 정도가 된다. 따라서 본 사례에 대한 종합복제정도를 위의 수식 (1)을 적용하면 다음과 같다.

종합복제 정도 = (10% × 0) + (40% × 0) + (5% × 0) + (10% × 0) + (15% × 0) + (10% × 32.7) + (5% × 32.5) + (5% × 5.4)

전체적인 종합복제 정도는 5.17%로서 결과적으로 매우 낮은 수치이며 따라서 프로그램의 복제나 도용은 매우 미약하거나 없었다고 판단할 수 있는 것이라 할 수 있다.

4. 결 론

본 연구에서는 소프트웨어 개발이 증대하면서 제기되고 있는 프로그램의 복제여부를 판단하는 내용이다. 본 연구는 실제 사례를 중심으로 프로그램의 복제정도 및 도용여부를 판단하기 위한 내용으로서 이를 위해 감정항목을 설정하고 감정 항목별 중요도를 설정하였으며 각 항목에 대한 복제정도를 산출하였다. 이를 바탕으로 하여 종합적인 복제정도 및 도용여부를 판단하였다.

참 고 문 헌

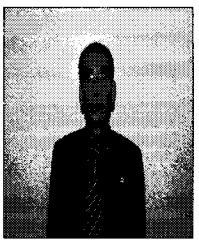
[1] 양사 프로그램.
[2] "프로그램 복제도 감정기법" 프로그램 심의조정위원회, 2002.
[3] J.Whitten. L.Bentley, System Analysis&Design Methods, IRWIN, 1994.
[4] T.Gildersleeve, Successful Data Processing Systems Analysis, 2nd edition, Englewood Cliffs, N.J. Prentice Hall, 1985.
[5] "유형별 소프트웨어 감정기법 표준화에 관한 연구" 프로그램 심의조정위원회, 2003.

김 용 원(Yoeng-Won Kim) [정회원]

- 1986년 고려대학교 수학과 졸업 (학사)
- 1988년 고려대학교 대학원 전산학과 졸업(석사)
- 1995년 고려대학교 대학원 전산학과 졸업(박사)
- 1995~현재 건양대학교 기업정보관리학과 교수

<관심분야>
데이터베이스 모델링, 데이터 검색, 데이터베이스 품질, 인터넷 데이터베이스 etc.

이 성 훈(Seong-Hoon Lee) [정회원]



- 1985년 한남대학교 전산공학과 졸업(학사)
- 1995년 고려대학교 대학원 전산학과 졸업(석사)
- 1998년 고려대학교 대학원 전산학과 졸업(박사)
- 1998~현재 백석대학교 정보통신학부 교수

<관심분야>
분산컴퓨팅, 이동통신, 소프트웨어 검증 etc.