Journal of the Ergonomics Society of Korea
Vol. 27, No. 3 pp.43-52, August 2008

철도사고 인적오류 분석을 위한 지원시스템 프레임웍 설계

신민주[1] · 백동현[2] · 김동산[3] · 윤완철[4]

[1]삼양 데이터 시스템 e-Biz 팀 / [2]한양대학교 경상대학 경영학부 / [3,4]KAIST 산업및시스템공학과

A Framework for Computerized Human Error Analysis System - Focused on the Railway Industry

Min Ju Shin[1], Dong Hyun Baek[2], Dong San Kim[3], Wan Chul Yoon[4]

[1]e-Biz Team, Samyang Data Systems Inc., Seoul, 110-725
[2]Department of Business Administration, Hanyang University, Ansan, 426-791
[3,4]Department of Industrial & Systems Engineering, KAIST, Daejeon, 305-701

ABSTRACT

Human errors are now considered as the most significant source of accidents or incidents in large-scale systems such as aircraft, vessels, railway, and nuclear power plants. As 61% of the train accidents in Korea railway involving collisions, derailments and fires were caused by human errors, there is a strong need for a systematic research that can help to prevent human errors. Although domestic railway operating companies use a variety of methods for analyzing human errors, there is much room for improvement. Especially, because most of them are based on written papers, there is a definite need for a well-developed computerized system supporting human error analyzing tasks. The purpose of this study is to propose a framework for a computerized human error analysis system focused on the railway industry on the basis of human error analysis mechanism. The proposed framework consists of human error analysis (HEA) module, similar accident tracking (SAT) module, cause factor recommendation (CFR) module, cause factor management (CFM) module, and statistics (ST) module.

Keyword: Human error analysis, Railway safety, Computerized system

1. 서 론

항공, 해양, 철도, 원자력 등과 같은 대형 시스템은 일상생활과 밀접하게 연관되어 사람들의 생활을 편리하게 하지만, 사고가 발생하게 되면 엄청난 인명과 재산상의 손해를 초래한다. 이러한 대형 시스템 사고의 대표적인 예로는 미국 TMI(Three Mile Island)원자력 발전소 사고(1979년), 영국의 킹스크로스 지하철역사 사고(1987년), 대구 지하철 화재 사고(2003년) 등으로 국가와 산업을 가리지 않고 발생하였다.

대형 시스템 사고의 원인 중 주요한 원인으로 손꼽히는 것이 인적오류(human error)이다. 해외에서는 인적오류의 원인 연구를 통해 해결방안을 모색하는 등 인적오류를 줄이

*본 연구는 국토해양부 미래철도기술개발사업의 연구비 지원(05철도안전B-02)에 의해 수행되었습니다..
교신저자: 백동현
주 소: 426-791 경기도 안산시 상록구 사1동, 전화: 010-8253-3307, E-mail: estarbaek@hanyang.ac.kr

기 위해 이미 많은 노력을 기울여 왔다. 그러나 국내에서는 인적오류의 중요성에 대한 인식은 늘고 있지만, 이에 대한 연구가 부족한 실정이다. 2006년 건설교통부가 발표한 철도안전종합계획을 보면 충돌, 탈선, 화재, 건널목 사고 등 열차 사고의 주요 원인으로 취급부주의(인적오류)가 61%로 가장 높게 나타났다. 이러한 결과는 국내 철도 산업에서도 인적오류 방지를 위한 대책 마련과 연구의 필요성이 증가하고 있음을 시사하고 있다. 그러나 현재까지 국내 철도 산업에 인적오류가 포함된 사고를 분석하기 위한 체계적인 방법이나 절차 등이 없으며, 인적오류와 관련된 원인별 항목도 제한적이어서 인적오류에 대한 철저한 원인 규명이 어려운 실정이다(김동산 등 2007).

인적오류 분석 사례는 유사 혹은 동종사고 분석의 해결을 위한 기준으로 활용될 수 있으나 국내의 경우 그러한 자료가 대부분 문서형식으로 되어 있고 데이터베이스화되어 있더라도 체계적인 분류체계에 의해 정리된 자료는 부재한 실정이다. 따라서 인적오류 분석체계를 전산화하여 효율적인 인적오류 분석을 지원하고, 적절한 정보를 담은 분석보고서의 자동 도출 등을 통해 분석자의 분석 노력을 절감시키고, 효과적인 인적행위 연구 및 대응방안 수립이 가능하도록 지원하는 전산시스템의 개발이 필요하다.

따라서 본 연구에서는 철도사고에 내재된 인적오류 분석을 위한 지원시스템의 프레임웍을 제시하고자 한다. 인적오류 분석을 위한 웹 기반 전산시스템이 구현된다면 사고 분석에 필요한 시간과 인력을 줄이고 체계적인 인적오류 분석이 가능할 것이다. 본 연구에서 제안한 프레임웍은 인적오류 분석 기능뿐만 아니라 유사사고 검색, 원인요소 추천 등 지능형 의사결정 지원 기능과 원인요소관리와 통계 기능 등으로 구성되어 있다.

본 논문은 다음과 같이 구성되어 있다. 먼저, 2장에서는 인적오류의 개념, 인적오류 분석을 위한 다양한 분석체계, 그리고 본 연구의 선행 연구를 통해 개발한 철도 인적오류 분석체계(HEAR)를 살펴보고, 본 연구에서 지능형 의사결정 지원을 위해 적용한 협업 필터링(collaborative filtering)과 연관규칙(association rule)에 대해 알아본다. 3장에서는 본 연구에서 제안한 인적오류 분석을 위한 지원시스템 프레임웍에 대해 설명한다. 마지막으로 4장에는 본 연구의 결론과 향후 과제에 대해 정리하였다.

2. 이론적 배경

2.1 인적오류의 개념

사람이 일으킬 수 있는 불안전한 행위(unsafe acts)는 그림 1과 같이 크게 오류(error)와 위반(violation)으로 나눌 수 있고, 오류는 일반적으로 실수(slip), 기억실패(lapse), 착오(mistake)의 3가지로 구분된다(Reason, 1990). 인적오류(human error)가 발생했다는 것은 의도한 목적을 이루기 위해 계획한 행위들이 실패하여 의도하지 않은 결과를 야기했다는 것을 의미한다.

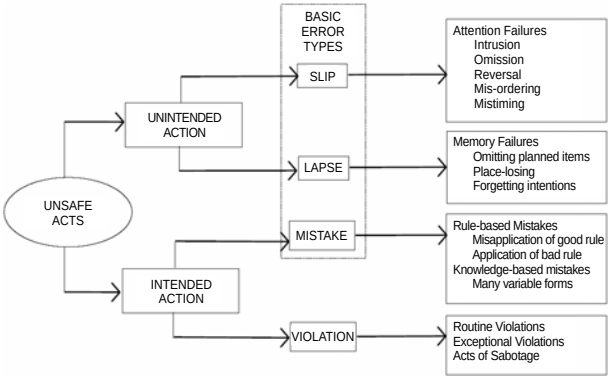

그림 1. Reason(1990)의 불안전한 행위 분류

계획 자체는 적절한데 행위가 계획대로 이루어지지 않은 경우를 '실수(slip)'라고 하고, 행위는 계획대로 이루어졌지만 계획이 부적절하여 실패한 경우를 '착오(mistake)'라고 한다. 기억의 문제로 생긴 오류를 따로 구분하여 '기억실패(lapse)'라고 하기도 한다. 위반은 안전하고 효율적인 작업을 위해 만들어진 규칙, 절차, 지시, 규정 등으로부터 벗어난 것을 의미한다.

실수는 주의 또는 인식의 실패와 관련된 것으로 대개 빨리 발견이 가능하여 피해나 손실로까지는 잘 이어지지 않지만, 기억실패는 실수보다 내면적이고 발견하기 어렵기 때문에 더 위험하다. 유지보수(maintenance) 작업과 같은 경우 작업을 끝마친 후 더 이상의 점검이 불필요하다고 생각할 때 문제가 더 악화될 수 있다. 여기서는 너무 늦어질 때까지 문제가 발견되지 않을 수 있기 때문이다. 착오는 기억실패보다 더 위험하다. 착오를 저지른 사람은 자신이 작업을 제대로 수행하고 있다고 생각하여 자신의 잘못을 알려주는 증거들이 종종 무시되기 때문이다.

2.2 인적오류 분석체계에 관한 연구

인적오류 분석체계와 관련된 연구는 국내외 산업 중 원자력 분야에서 가장 먼저 시작되어 다양한 분석 방법과 도구들이 개발되었고, 항공 분야에서는 인적오류 분석체계 및 방법론을 중심으로 연구가 활발히 진행되고 있다. 철도 분야는 인적오류 분석에 대한 연구를 진행한 지 얼마 되지 않

아 관련 연구가 부족한 편이지만, 영국을 중심으로 활발한 연구가 진행되고 있다(KAIST, 2007).

김동산 등(2007)의 연구에서는 철도, 항공, 원자력, 기타 분야로 산업을 분류하여 해당 산업에서 사용하거나 특정 산업에서의 활용을 목적으로 개발된 분석기법들을 조사하였고, 각 분석기법에 대한 장단점을 도출하였다. 대표적인 인적오류 분석체계로는 원자력 분야의 HPES(Human Performance Enhancement System)와 HPIP(Human Performance Investigation Process), 항공 분야의 HFACS (Human Factors Analysis and Classification System)와 TRACEr(Technique for the Retrospective Analysis of Cognitive Errors), 그리고 기타로는 CREAM(Cognitive Reliability and Error Analysis Method)와 HFIT(Human Factors Investigation Tool) 등이 있다. 또한, 김동산 등(2007)은 대표적인 인적오류 분석기법들의 장단점을 토대로 국내 철도 산업에의 적용 방안을 제시하였다. 사고 분석을 위한 전산시스템으로는 TapRooT®, Apollo, ProAct®, Reason® 등이 있는데 '인적오류'의 분석에 특화된 전산시스템은 거의 없고 여러 측면에서 아직 개선되어야 할 사항이 많다.

본 연구의 선행 연구에서는 국내 철도 산업에 적합한 인적오류 분석체계(HEAR: Human Error Analysis and Reduction)를 개발하였다(KAIST, 2006, 2007). HEAR는 국내 철도 분야의 분석 절차 및 방법에 대한 조사와 국내외 산업계(원자력, 항공, 철도 등)의 인적오류 분석 절차 및 기법에 대한 동향, 그리고 인적오류 분석체계를 설계하는데 있어서 이론적 토대가 되는 사고 인과관계 모형과 인간 의사결정 모형을 참고하여 만들어졌다.

그림 2는 HEAR의 분석 절차 및 정보 흐름을 도식화한 것이다. HEAR는 사고개요 파악에서부터 기본상황분석, 원인요소분석 그리고 예방대책평가까지 총 6단계, 17개 세부단계로 구성되어 있으며, 각 단계별로 분석에 필요한 분석도구와 각 단계별 산출물이 정의되어 있다. 각 단계별 분석 양식과 절차에 대한 자세한 사항은 KAIST(2007) 연구보고서를 참조하기 바란다.

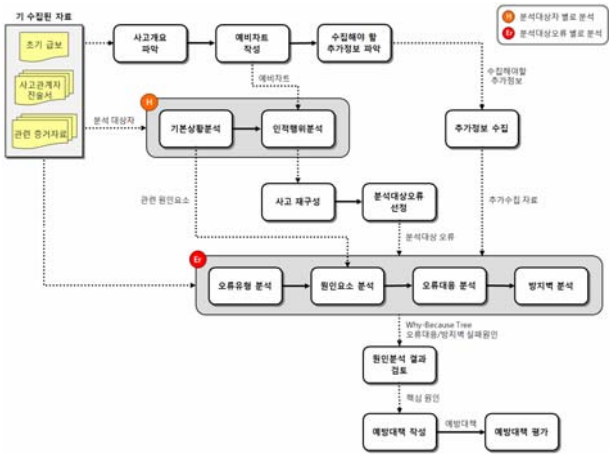

그림 2. HEAR 분석절차 및 정보흐름

2.3 협업 필터링(collaborative filtering)

협업 필터링은 사용자의 선호도와 관심 표현을 바탕으로 비슷한 패턴을 가진 사용자들을 식별하여 사용자의 취향이나 생활 형태, 유사성에 따라 관련 아이템을 추천하는 서비스를 제공하기 위해 사용되며, 주로 인터넷 쇼핑몰에서 이용되고 있는 성공적인 상품추천기법 중의 하나이다(Sarwar et al., 2000).

협업 필터링은 크게 사용자 기반 협업 필터링과 아이템 기반 협업 필터링으로 나뉘는데(Herlocker et al., 1999), 사용자 기반 협업 필터링은 사용자 간의 연관관계를 파악한 후 이를 기반으로 추천하는 알고리즘이다. 따라서 사용자 집단의 선호도가 비슷할수록 상대적으로 높은 예측력을 보이고 사용자 집단의 선호도가 상이할수록 예측의 정확성이 떨어진다는 한계점을 가지고 있다. 아이템 기반 협업 필터링은 아이템 간의 유사성, 즉 고객이 선호도 등급을 입력한 기존 상품들과 추천하고자 하는 상품들 간의 유사성을 측정하여 어떤 특정 고객이 어떤 상품을 선호하는지 예측하여 추천하는 방식이다(홍태호 등, 2005). 아이템 기반 협업 필터링의 한계점은 아이템에 대한 사용자들의 선호도를 사용하므로 사용자 간의 선호도가 상이할 경우 예측력이 떨어질 수 있다는 것이다.

협업 필터링 기반의 추천 과정은 다음과 같다. 먼저, 추천 대상과 유사한 이웃을 찾는다. 추천 대상과 유사한 이웃을 찾는 방법으로는 피어슨 상관계수(Pearson correlation), 코사인(cosine) 등의 방법이 있다. 피어슨 상관계수는 사용자 간의 유사도 가중치를 계산하는데 사용되는 방법으로 추천의 정확도가 높아 널리 사용되고 있다 (Sarwar et al., 2001). 두 사용자 사이의 유사도를 r이라는 기호로 표기하고, 그 값은 −1에서 1까지의 값을 취한다. 상관계수의 부호가 양(+)일 경우에는 양의 상관관계가, 음(−)일 경우에는 음의 상관관계가 있는 것이고 부호에 상관없이 1에 가까울수록 상관이 강하다는 것을 의미한다. 또한 상관관계가 존재하지 않을 경우에는 상관계수의 값이 0에 가깝다.

변수 x, y의 상관계수 r을 계산하는 방법은 다음과 같다.

$$r(S) = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2 \sum_{i=1}^{n}(y_i - \bar{y})^2}} \quad \cdots (1)$$

또 다른 방법인 코사인은 아이템과 아이템 간의 관련성을 근거로 추천해주는 방법으로 계산하는 식은 (2)와 같다.

$$Cos(\vec{x}, \vec{y}) = \frac{\vec{x} \cdot \vec{y}}{\| \vec{x} \|_2 \| \vec{y} \|_2} \quad \cdots (2)$$

위와 같은 방법으로 이웃을 찾은 후에는 예측에 사용될 이웃의 수를 결정하기 위하여 Thresholding 기법과 Best-n-neighborhood 기법을 사용한다. Thresholding 기법은 예측된 값 중 일정 값 이상의 대상만을 사용하도록 제안하는 것이고, Best-n-neighborhood 기법은 n 명의 대상만을 사용해서 예측하도록 제안하는 방법이다(이기현 등, 2002; Resnick et al., 1994).

협업 필터링의 대표적인 연구로 GroupLens(Resnick et al., 1994)는 Netnews와 영화에 대해 협업 필터링을 사용하였다. Netnews를 사용하는 고객들에게 새로운 책, 요리법, 음악, 영화 등의 기사에 대한 평가를 받아 이를 기반으로 목표고객과 다른 고객 간의 유사도를 계산한 후 최근접 이웃을 구성하였다. 웹 기반의 협업 필터링에 의한 추천 시스템으로는 Ringo, Video Recommender(Sarwar et al., 2000), 화장품을 도메인으로 한 가상쇼핑몰의 고객 추천시스템(김재경 등 2005) 등이 있다.

2.4 연관규칙(association rule)

데이터 마이닝(data mining)은 방대한 양의 데이터 중에서 의미 있는 패턴이나 규칙을 찾기 위한 일련의 분석 과정으로 주로 인터넷 상거래 사이트에서 고객 맞춤 서비스를 제공하는 목적으로 사용되고 있다. 그 중에서 연관규칙은 인터넷 쇼핑몰 고객들의 장바구니 분석, 고객 정보와 구매 패턴 간의 연관성 파악 등에 활용된다.

연관규칙은 다음과 같은 형태를 취하게 된다.

(Item 집합 X) → (Item 집합 Y) … (3)
: If X then Y, 만일 X가 일어나면 Y가 일어난다.

식 (3)은 어떤 Item 집합 X의 존재가 다른 Item 집합 B의 존재를 암시하는 것을 의미한다. 이것은 함께 구매하는 상품의 조합이나 서비스의 패턴을 발견하거나 특정 제품이나 사건들이 동시에 발생하는 패턴을 파악하는데 이용할 수 있다. 이러한 연관규칙을 인적오류 분석에 응용하면, 원인요소 X가 특정 인적오류의 원인이면 어느 정도의 확률을 가지고 Y 역시 그 인적오류의 원인이 될 수 있음을 의미한다.

일반적으로 연관규칙을 계량화할 때 참고하는 지표로는 지지도(support), 신뢰도(confidence), 향상도(lift) 등 세 가지가 있는데, 이를 인적오류 원인분석에 응용한 것은 표 1과 같다.

표 1. 연관규칙의 측정 지표

지표	설명	도출식
지지도 (Support)	원인요소 A, B가 사고를 유발시킨 오류에 동시 존재할 수 있을 확률로 정의	$S = P(A \cap B) = \frac{\text{원인요소 } A\text{와 } B\text{를 포함하는 오류}}{\text{전체 오류수}(N)}$
신뢰도 (Confidence)	원인요소 A를 포함하는 오류 중에서 원인요소 B가 포함될 확률로 정의	$C = P(B\|A) = \frac{P(A \cap B)}{P(A)} = \frac{\text{원인요소 } A\text{와 } B\text{를 포함하는 오류}}{\text{원인요소 } A\text{를 포함하는 오류수}}$
향상도 (Lift)	원인요소 A를 포함한 경우, 오류에 원인요소 B를 포함하는 경우와 원인요소 B가 임의로 포함되는 경우의 비를 측정하는 지표	$L = \frac{P(B\|A)}{P(B)} = \frac{P(A \cap B)}{P(A)P(B)}$

지지도는 오류 트랜잭션 내에 원인요소 A, B가 모두 포함된 트랜잭션의 비율을 나타냄으로써 원인요소 사이에 연관 정도의 중요도를 나타낸다. 신뢰도는 원인요소 A를 포함하고 있는 트랜잭션 중 원인요소 B를 포함하고 있는 트랜잭션의 비율을 나타냄으로써 아이템들 사이에 연관 정도를 나타낸다. 마지막으로 향상도는 원인요소 A가 포함된 오류의 경우 해당 오류에 원인요소 B를 포함하는 경우와 원인요소 B가 임의로 포함되는 경우의 비를 나타낸 것으로 향상도의 값이 1 보다 크면 양의 상관관계를, 1 보다 작으면 음의 상관관계의 의미를 갖는다. 이 지표들을 활용하는 방법은 먼저 사용자가 정의한 최소 지지도 값을 넘는 규칙들만 골라낸 후에 그 중 최소 신뢰도 값을 넘는 규칙들을 골라내면 된다. 향상도 값은 연관규칙에 연계되어 있는 원인요소가 많은 인적오류의 분석에 활용하면 더 의미 있는 규칙을 추출할 수 있다.

3. 인적오류 분석을 위한 지원시스템 프레임웍

본 장에서는 철도 산업에서 인적오류가 내재된 사고의 분석을 위한 지원시스템(CAS-HEAR: Computer Aided System for Human Error Analysis and Reduction, 이하 CAS-HEAR)의 프레임웍에 대해 설명한다. CAS-HEAR는 인적오류 분석 절차 및 양식, 관련 정보 제공, 통계 기능 등을 실시간으로 제공함으로써 분석 결과의 질을 높이고, 분석자의 인적오류 분석에 대한 의사결정을 효과적이고 지

능적으로 지원하는 인적오류 분석 지원시스템이다. CAS-HEAR는 시스템의 사용자인 분석자가 분석 절차에 대해서 암기해야 할 필요성을 줄여주고 분석과 보고서 작성에 소요되는 인력과 시간을 절감시켜 준다.

그림 3은 CAS-HEAR의 프레임웍으로, CAS-HEAR는 인적오류 분석 모듈, 유사사고 검색 모듈, 원인요소 추천 모듈, 원인요소 관리 모듈, 통계 모듈, 그리고 인증관리 모듈 등 6개의 모듈로 구성되어 있으며, 사고 분석자나 관리자 또는 일반 사용자가 내·외부 네트워크를 통해 CAS-HEAR에 접속한다. CAS-HEAR는 필요시 국토해양부, 교통안전공단, 철도운영자 등의 정보시스템과 데이터를 주고받을 수 있다.

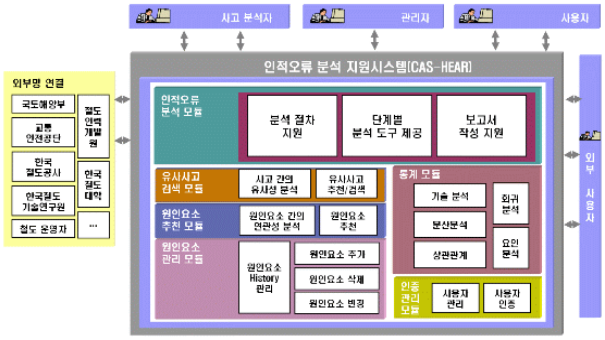

그림 3. CAS-HEAR 프레임웍

인적오류 분석 모듈은 인적오류 분석을 위한 절차 및 도구, 그리고 보고서 작성을 지원해 주며, 유사사고 검색 모듈은 현재 분석 중인 사고와 유사한 과거사고 사례를 검색하여 제공함으로써 분석자가 보다 효과적으로 사고를 분석할 수 있도록 한다. 원인요소 추천 모듈은 원인요소 간의 연관성을 분석하여 인적오류의 원인을 분석하는 과정에서 사용자에게 의미 있는 원인요소를 추천해줌으로써 분석을 지원한다. 통계 모듈은 분석된 사고들을 대상으로 시기, 장소, 분석 대상자(기관사, 역무원 등)별 사고발생 빈도, 상황요소별 인적오류 빈도, 원인요소 간의 상관관계, 오류유형별 발생 빈도 등을 통계 처리하여 분석 도중 필요한 자료가 있을 경우 실시간으로 제공할 수 있으며 분석 결과에 대한 검색, 조회가 용이하다. 이외에 원인요소 관리 모듈은 각 원인요소 항목에 대한 추가, 삭제, 변경 등의 기능을 담당하며 인증관리 모듈은 사용자 관리 및 인증을 담당한다.

3.1 인적오류 분석 모듈

인적오류 분석 모듈은 인적오류 분석을 위한 분석 절차 및 단계별 분석 도구를 제공하여 분석자가 순차적으로 분석을 진행함으로써 사고를 유발시킨 인적오류와 인적오류의 발생 원인이 되는 원인요소에 대한 분석이 이루어지도록 지원한다.

그림 4와 같이 분석 준비, 정보수집, 사고 재구성, 원인분석, 예방대책 마련, 분석 결과 보고 단계로 구성되며, 각 단계별로 필요한 정보는 DB로부터 제공받으며 모든 분석 결과는 분석 DB에 저장된다. 또한, 분석 DB에 저장된 분석 결과는 분석 결과 보고 단계에서 자동으로 보고서를 생성할 수 있도록 지원하고, 분석자가 요청할 경우 조회가 가능하다.

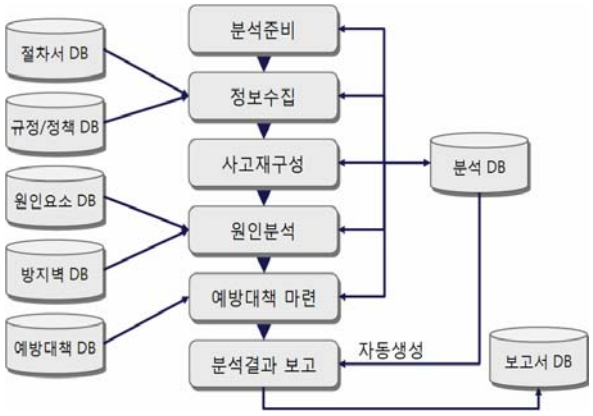

그림 4. 인적오류 분석 모듈의 사고분석 절차

본 연구를 통해 구현된 인적오류 분석 모듈은 현재 웹 환경에서 프로토타입 수준으로 개발되어 있으며, 이 시스템에 대한 자세한 내용은 신민주(2008)를 참조하기 바란다. 개발된 프로토타입 시스템을 현업에 실제 적용하기 위해서는 추가 개발이 필요하다.

3.2 유사사고 검색 모듈

유사사고 검색 모듈은 인적오류 분석 시에 효과적인 사고분석이 가능하도록 분석자를 지원하는 기능으로, 분석 중인 사고와 유사한 사고를 과거사고 사례로부터 검색해준다.

그림 5는 유사사고 검색 모듈에 대한 논리적 모델을 도식화한 것이다. 먼저, 분석자가 대상 사고를 분석하는 중간에 시스템에게 유사사고 검색을 요청하면 시스템은 사고 사례 데이터베이스를 검색하여 대상 사고와 유사한 사례를 찾는다. 본 연구에서는 유사사고 검색을 위해 인적오류 분석 모듈의 한 단계인 '기본상황분석' 결과를 토대로 피어슨 상관계수를 이용하여 유사 사례를 찾는다. 시스템은 Pearson 상관계수의 결과 값을 토대로 유사사고 리스트를 작성하여 분석자에게 제공하게 되고, 분석자는 시스템이 제공하는 추

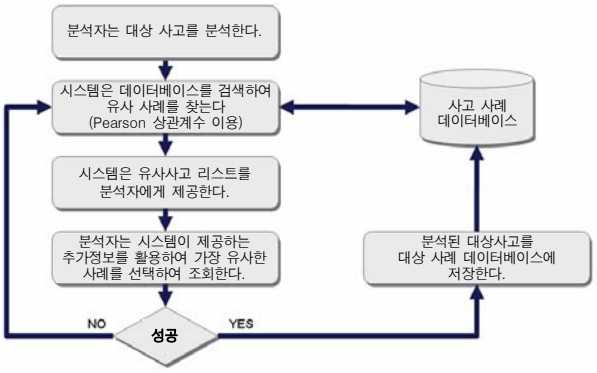

그림 5. 유사사고 검색 모듈 논리적 모델

가 정보를 활용하여 가장 유사한 사례를 선택, 조회할 수 있다. 시스템이 추천해 준 유사사고 리스트가 사용자의 마음에 들었다면, 시스템은 검색 결과를 성공으로 판단하여 유사사고 검색 모듈은 종료된다. 분석 중인 사고는 분석이 종료되면 사고 사례 데이터베이스에 저장된다. 반면에, 사용자가 시스템의 유사사고 추천 리스트가 마음에 들지 않았다면 시스템은 데이터베이스를 다시 검색하여 유사 사례를 찾게 된다.

유사사고 검색 모듈을 설명하기 위해 국내 철도사고 사례 중 인적오류로 인한 사고 사례를 분석대상 사고로 적용해 보았다. 분석대상 사고는 대구선 고모역 구내 관광열차 탈선 사고로 2006년 2월 12일 오전 5시 13분경 강릉역을 떠나 부산역으로 가던 제4513 금강산 관광열차(강릉 22:45 → 부산 07:56)가 대구선 고모역 접근 시 운전취급 계획을 변경하였고, 관제사가 사전 통보 없이 진로를 변경하여 선로전환기가 전환되면서 기관차 및 객차 2량이 궤도에서 이탈된 사고였다.

기본상황분석은 사고 당시의 기본적인 상황요소들에 대해 분석하는 것으로 작업자, 직무, 작업환경, 그리고 조직 관련 상황분석표를 이용하여 분석하는데, 그림 6은 대구선 고모역 구내 관광열차 탈선 사고에 대한 작업자 관련 기본상황분석을 수행한 결과를 보여준다. 직무, 작업환경, 그리고 조직 관련 기본상황분석도 이러한 형태의 상황분석표를 이용하여 분석한다. 분석자는 상황분석표에 있는 각 항목들의 내용을 분석하고 사고와의 연관성을 5단계로 평가하도록 되어 있다(매우 낮음: 1, 낮음: 2, 보통: 3, 높음: 4, 매우 높음: 5).

분석대상 사고와의 유사성을 비교하기 위해 임의로 데이터를 작성하여 가상 사고 사례 A, B, C, D, E를 마련하였다. 표 2는 분석대상 사고와 사고 사례 A, B, C, D, E의 기본상황분석(작업자, 직무, 작업환경, 조직)의 결과인 총 45개의 항목에 대해 사고와의 연관성을 수치화한 결과이다. 각 항목에 대해 1부터 5까지 숫자로 표시가 되어 있는데, 이는

항 목	내 용	사고와의 연관성					비고
		매우 낮음	낮음	보통	높음	매우 높음	
1. 해당직무 경력	(1) 년 (8) 개월, 전체 근무 경력				√		
2. 신체검사	1) 최근 검사일: ( ) 년 ( ) 월 2) 신체검사 결과 특이사항	√					
3. 적성검사	1) 최근 검사일: (1997) 년 (3) 월 2) 적성검사 결과 특이사항		√				
4. 관련 규정 및 절차 숙지여부	□ 모두 숙지 □ 일부 숙지하지 않음 ■ 잘못 이해하고 있었음					√	
5. 교육/훈련	1) 최근 교육/훈련받은 날짜 (2005) 년 (10) 월 (10, 12) 일 2) 교육/훈련 시간: (4) 시간 3) 교육/훈련 내용: 신설 대구선 구간 운전 취급 요령					√	
6. 질병/술	1) 사고 당시 앓고 있던 질병: (의 족) 2) 취중 작업 여부: □ 예 □ 아니오		√				
7. 수면시간	1) 사고 전날 수면시간: (4) 시간 2) 사고 전 1주일 평균 수면시간: (8) 시간			√			
8. 당시 감정상태	■ 보통 □ 긴장 □ 우울 □ 흥분 □ 기타( )		√				
9. 직무 및 역할에 대한 만족도	■ 높음 □ 보통 □ 낮음		√				
10. 사고 이력	1) ( ) 년 ( ) 월 (없음) 사고 사고원인: ( ) 2) ( ) 년 ( ) 월 (없음) 사고 사고원인: ( )	√					

그림 6. 분석대상 사고의 작업자 관련 상황분석표

표 2. 분석대상 사고와 사고사례의 데이터 형태

항 목	분석대상 사고	사례 A	사례 B	사례 C	사례 D	사례 E
해당직무 경력	4	2	1	2	1	4
신체검사	1	5	2	3	3	2
적성검사	2	4	3	5	4	3
관련 규정 및 절차 숙지여부	5	3	4	3	3	4
교육/훈련	5	2	5	3	4	4
질병/술	2	4	3	4	2	3
⋮	⋮	⋮	⋮	⋮	⋮	⋮
작업자 교육 및 훈련	4	2	4	2	4	5
시스템 설계 및 유지보수	4	2	1	2	5	5
작업자간 의사소통 및 지침의 전달	5	1	3	2	2	5
업무 분담, 책임 및 권한 부여	5	3	3	1	3	5
포상, 승진 및 처벌 규정	2	4	4	5	4	3
조직에 대한 작업자들의 전반적인 만족도	1	4	2	4	2	2

각 항목이 해당 사고가 발생하는데 얼마만큼의 연관성을 가지고 있는지를 나타낸다(1: 매우 낮음, 5: 매우 높음). 예를 들어 분석대상 사고에서는 관련 규정 및 절차 숙지여부, 교육/훈련, 작업자간의 의사소통 및 지침의 전달, 업무 분담, 책임 및 권한 부여는 해당 항목에 대한 값이 5로 사고와의 연관성이 매우 높은 것으로 나타났음을 알 수 있다.

표 3. 분석대상 사고와 사고 사례 간의 유사도 검증

	사례 A	사례 B	사례 C	사례 D	사례 E
분석 대상 사고	−.610**	.307*	−.605**	.256	**.824****

*. Correlation is significant at the 0.05 level(2−tailed)
**. Correlation is significant at the 0.01 level(2−tailed)

사례 검증 결과를 보면, 분석대상 사고와 사례 A, C는 음의 상관관계(−.610, −.605)가 있다. 분석대상 사고와 사례 B, D는 상관관계가 낮은 편이지만(.307, .256), 사례 E는 매우 강한 양의 상관관계(.824)를 가지고 있으며 이는 통계적으로도 매우 유의하다. 기본상황분석 결과를 비교한 결과 분석대상 사고와 사례 E는 강한 양의 상관관계를 갖고 있다는 것을 알 수 있다.

통계 분석의 경우 데이터에 내포된 의미를 전부 반영하지 못하는 경우가 종종 있다. 따라서 직무, 작업환경, 조직, 작업자 등 각 분야별로 상관분석을 실시하였고 그 결과는 표 4에 정리되어 있다. 표 4는 사고의 종류, 분석대상 사고와 사고사례 간의 전체 유사도 및 분야별 유사도를 보여준다. 시스템이 이와 같은 결과를 제공하면 분석자는 사고의 종류와 유사도 값을 참고하여 유사 사례를 선택적으로 조회할 수 있다.

표 3과 같이 기본상황분석 전체 결과를 가지고 분석했을 때는 사례 E만 유의한 값을 가졌으나, 표 4의 분야별 유사도에서 직무 분야만을 가지고 분석을 했을 때에는 사례 D도 유의한 상관성이 있는 것으로 나타났다. 즉, 분석대상 사고와 사례 D는 사고의 종류가 각각 탈선 사고와 건널목 사고로 상이하고 전체 유사도도 0.256으로 낮은 편이지만 사고가 발생한 직무 관련 상황만을 고려한 경우 유사한 측면이 있음을 알 수 있다. 또한 작업환경 분야의 분석 결과에서는 사례 B가 분석대상 사고와 유의한 양의 상관관계가 있음을 알 수 있다. 이와 같이 기본상황분석 전체 및 세부 분야별 유사도를 함께 고려하는 것은 과거 사례로부터 유사한 사례를 검색하는데 유용한 정보를 제공한다.

표 4. 분석대상 사고와 사고 사례 간의 유사도 검증 결과

분석대상 사고의 사고 종류: **탈선 사고**

	사고 종류	전체 유사도	분야별 유사도			
			작업자	직무	작업환경	조직
사례 A	직무사상 사고	−.610**	−.571	−.250	−.564*	−.737*
사례 B	시설물 사고	.307*	.474	.066	**.558***	.189
사례 C	탈선 사고	−.605**	−.517	−.354	−.645**	−.819**
사례 D	건널목 사고	.256	−.202	**.611***	.153	.311
사례 E	탈선 사고	**.824****	**.749***	**.773****	**.867****	**.951****

*. Correlation is significant at the 0.05 level (2−tailed)
**. Correlation is significant at the 0.01 level (2−tailed)

3.3 원인요소 추천 모듈

인적오류 분석 모듈의 원인분석 단계(그림 4 참조)는 인적오류의 원인을 심층적으로 분석하는 단계인데, 이를 위해 분석자는 시스템에서 제공하는 13개 카테고리, 총 149개의 원인요소 항목들 중에서 분석대상 인적오류의 원인이 되는 항목들을 일일이 선택해야 한다. 일반적으로 하나의 인적오류에 대해 여러 개의 원인요소가 관련되어 있음을 감안할 때 이렇게 많은 원인요소 항목으로부터 분석자가 일일이 적합한 항목을 찾는 것은 효율적이지 않다. 원인요소 추천 모듈은 원인요소 항목들 간의 연관성을 미리 분석하여 인적오류의 원인요소를 밝혀내는 과정에서 분석자에게 적합한 원인요소를 추천해 주는 기능을 담당한다.

원인요소 추천은 연관규칙의 지지도와 신뢰도 값을 이용하는데, 지지도의 값이 일정 수준 이상이면서 신뢰도의 값이 일정 수준 이상인 규칙을 추천해주는 방식이다. 트랜잭션의 수가 많을 경우 신뢰도 값을 대신해서 향상도 값을 활용하는 것이 의미 있는 규칙을 발견하는데 유리하다.

원인요소 추천 모듈을 설명하기 위해 가상으로 5개의 오류 원인분석 결과를 준비하였다. 표 5는 과거에 발생한 인적오류의 원인분석 결과로부터 도출된 것으로, 인적오류별로 어떠한 원인요소가 관련되었었는지를 표로 정리한 것이다. 예를 들어, 오류 A는 원인요소 a, c, d, e, f로 인해 발생했음을 보여주고 있다.

표 5와 같은 인적오류 원인분석 결과가 데이터베이스에 저장되어 있는 경우, 원인요소 추천 모듈은 분석자가 분석 중인 오류에 대한 원인으로 특정 원인요소를 선택하였을 때 그 원인요소와 연관성이 높은 다른 원인요소를 추천해 준다. 예를 들어, 분석자가 현재 분석 중인 오류의 원인요소로 a를 선택한 경우, 시스템은 원인요소 a와 다른 원인요소 간의 연관성 정도를 계산하여 가장 연관성이 높은 원인요소를 추천해 준다.

표 6은 원인요소 a와 다른 원인요소 간의 지지도, 신뢰도, 향상도 값을 표 1의 도출식을 이용해 계산한 결과이다. 여기서는 사용된 오류의 수가 적으므로 지지도와 신뢰도만을 사용하여 연관규칙을 도출하였다.

표 5. 인적오류 원인분석 결과

	원인 요소 a	원인 요소 b	원인 요소 c	원인 요소 d	원인 요소 e	원인 요소 f	원인 요소 g	원인 요소 h
오류 A	√		√	√	√	√		
오류 B	√		√	√	√		√	
오류 C	√	√		√	√	√	√	√
오류 D		√		√		√		
오류 E		√		√			√	√

표 6. 원인요소 a와 다른 원인요소 간의 연관성 결과

원인요소 a	원인 요소 b	원인 요소 c	원인 요소 d	원인 요소 e	원인 요소 f	원인 요소 g	원인 요소 h
지지도 (Support)	.2	.4	.6	.6	.4	.4	.2
신뢰도 (Confidence)	.33	.66	1	1	.66	.66	.33
향상도 (Lift)	.11	.33	1	1	.33	.22	.16

계산된 결과를 토대로 지지도의 값이 일정 수준 이상(예, 0.4 이상)인 규칙들을 찾아낸 후, 2차 선별 작업으로 신뢰도 값이 일정 수준 이상(예, 0.7 이상)인 연관규칙을 찾는다. 그 결과 원인요소 d와 e가 원인요소 a와 연관성이 높은 것으로 판정되어 분석자에게 분석 중인 인적오류의 원인요소로 추천하게 된다. 지지도와 신뢰도 수준의 선별 기준은 사례의 수와 시스템 상태에 따라 달라질 수 있어 기준 값 설정을 위해서는 일정 기간의 시행착오(trial and error)가 필요하다.

3.4 원인요소 관리 모듈

원인요소 관리 모듈은 원인요소 항목을 관리하는 기능으로 사고분석 결과를 바탕으로 원인요소 항목의 추가, 삭제, 수정 등을 지원하여 그 결과를 원인요소 데이터베이스에 저장한 후 새로운 사고 분석 시에 사용된다. 또한, 원인요소 항목의 추가, 삭제, 수정 등에 따른 항목별 History를 관리하고 있다. 이렇게 함으로써 원인요소 항목에 유연성 및 시계열성을 반영하여 기존의 원인요소 데이터베이스 내용을 지속적으로 관리해줄 수 있으며 시기별 원인요소 발생 빈도 조회, 발생 가능한 원인요소를 유추하여 그와 관련된 사고 발생을 예방할 수 있다.

그림 7은 원인요소 관리 모듈의 논리적 모델을 도식화한 것이다. 원인요소 관리 모듈은 크게 4가지 기능을 가지고 있는데, 첫 번째로 원인요소 추가 기능은 새로운 원인요소를 데이터베이스에 추가함으로써 원인요소가 추가된 이후부터는 분석자가 사고 분석 시에 추가된 원인요소를 사용할 수 있도록 한다. 두 번째로, 원인요소 삭제 기능은 기존에 저장되어 있던 원인요소 중 일부 항목을 삭제하는 기능으로 더 이상 사용하지 않는 원인요소를 데이터베이스부터 삭제하는 기능이다. 세 번째로 원인요소 수정 기능은 기존의 원인요소에 대한 이름, 설명 등의 정보를 새로운 내용으로 변경하는 기능으로 시간의 흐름을 반영하여 원인요소의 정보를 업데이트 할 수 있도록 한다. 마지막 기능은 원인요소 History 관리 기능으로 각각의 원인요소에 대해 추가, 삭제, 변경된 사항들에 대한 과거 이력을 관리함으로써, 원인요소 항목의 변화 추이를 통한 예방 교육을 강화할 수 있다. 이때 시스템 관리자는 사고 DB의 정보를 이용한 통계 모듈의 결과를 제공받게 되고, 그 결과를 토대로 원인요소를 추가, 삭제, 변경, 관리한다.

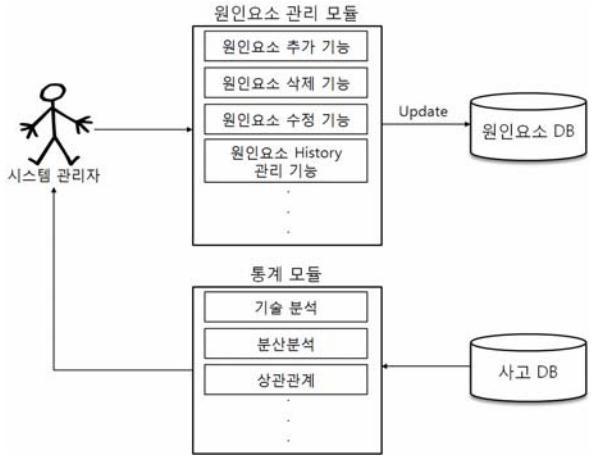

그림 7. 원인요소 관리 모듈의 논리적 모델

3.5 통계 모듈

통계 모듈은 분석된 사고들의 정보를 제공받아 여러 가지 의미 있는 범주로 사고의 원인을 통계 처리하여 인적오류의 경향을 파악하고 분석자, 관리자, 외부 이용자에게 유용한 통계 정보를 제공하여 효율적인 인적행위 연구 및 대책안 수립이 가능하도록 지원한다.

표 7은 통계 모듈에서 지원 가능한 분석기법들에 대한 설명과 해당 분석기법이 어떻게 활용될 수 있는지 예를 들어 설명한 것이다.

앞에서 제시한 분석 기법을 이용하면 그림 8과 같은 통계 지원 결과를 얻을 수 있는데, 자료는 2002년에 발행된 도시철도사고사례집을 토대로 기술 분석을 실시한 결과이다. 년도와 원인별 사상 사고 발생현황을 그래프로 나타낸 것으

표 7. 통계 모듈에서 지원하는 분석기법

분석 기법	설 명	적용 예
기술 분석	변수 값들이 이루는 분포의 특성을 알아보는데 이용 -평균값/최빈값/대표값/표준편차/분산/최저값/최고값	• 년도, 월, 요일, 시간, 장소별 사고 발생 빈도 • 상황요소별 인적오류 발생 빈도 • 오류유형별 발생 빈도
분산 분석	두 집단 이상의 평균 간의 차이를 검증하는데 이용	• 분석 대상자(예, 운영사별 기관사) 간의 사고 발생률의 차이
상관 관계	하나의 변수가 다른 변수와 어느 정도 밀접한 관련성을 갖고 변화하는가를 알아보기 위해 이용	• 사고와 원인요소 간의 상관관계 • 사고와 분석 대상자 간의 상관관계
회귀 분석	변수들 간의 상호관계를 분석하고 특정 변수(독립 변수)의 변화로부터 다른 변수(종속변수)의 변화를 예측하기 위해 이용	• 사고에 영향을 미치는 요인 규명
요인 분석	변수들 간의 상관관계를 이용하여 서로 유사한 변수들끼리 묶어주는 방법	• 원인요소 항목에 대한 요인 분석

로 x축은 사상 사고에 대한 원인, y축은 발생 건수, z축은 년도를 의미한다. 이러한 통계자료는 사용자의 편의에 따라 표 또는 여러 형태의 그래프로 변환할 수 있다.

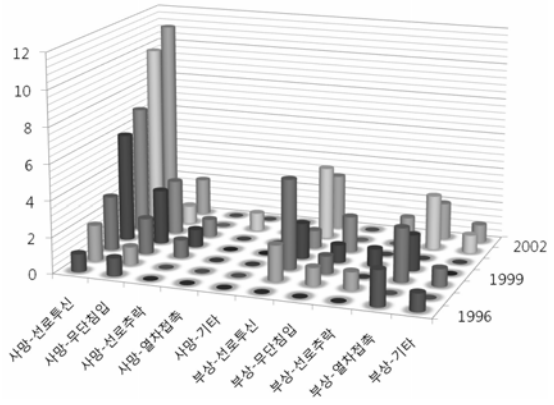

그림 8. 통계 지원 화면의 예(기술 분석)

4. 결론 및 검토

본 연구에서는 철도사고에 내재된 인적오류 분석을 위한 지원시스템의 프레임웍을 제시하였다. 제안된 지원시스템 프레임웍은 인적오류 분석 기능뿐만 아니라 유사사고 검색, 원인요소 추천 등 지능형 의사결정 지원 기능과 원인요소관리와 통계 기능 등으로 구성되어 있다. 인적오류 분석 모듈은 인적오류 분석을 위한 절차 및 도구, 그리고 보고서 작성을 지원해 주며, 유사사고 검색 모듈은 현재 분석 중인 사고와 유사한 사고를 과거사고 사례로부터 검색하여 제공함으로써 분석자가 보다 효과적으로 사고를 분석하도록 한다. 원인요소 추천 모듈은 원인요소 간의 연관성을 분석하여 인적오류의 원인을 분석하는 과정에서 사용자에게 의미있는 원인요소를 추천해줌으로써 분석을 지원한다. 통계 모듈은 분석된 사고들을 대상으로 시기, 장소, 분석 대상자(기관사, 역무원 등)별 사고발생 빈도, 상황요소별 인적오류 빈도, 원인요소 간의 상관관계, 오류유형별 발생 빈도 등을 통계 처리하여 분석 도중 필요한 자료가 있을 경우 실시간으로 자료를 제공할 수 있으며 분석 결과에 대한 검색, 조회가 용이하다. 이외에 원인요소 관리 모듈은 각각의 원인요소 항목에 대한 추가, 삭제 등의 기능을 담당하며 인증관리 모듈은 사용자 관리 및 인증을 담당한다.

본 연구의 결과는 국토해양부, 한국철도공사 등과 같은 철도사고를 조사 · 분석하는 기관에서도 인적오류 분석을 위한 지원시스템을 개발하고자 할 때에 활용할 수 있다. 또한 타 산업 분야에서도 인적오류 분석 지원시스템을 구축할 때 분석자에게 필요한 분석 및 의사결정 지원 기능을 설계하는데 활용할 수 있다.

인적오류 분석을 위한 지원시스템의 프레임웍을 제시함에 있어 다음과 같은 연구의 한계점과 향후 연구방향을 제안한다. 첫째, 연구의 한계점으로 유사사고 검색 모듈의 경우, 기본상황분석 결과를 값으로 변환하여 사용하는데 각각의 항목들 간의 가중치를 동일하게 계산한다는 한계점이 있다. 따라서 유사사고 모듈에서 사용하는 기본상황분석의 항목별 가중치를 다르게 부여하는 연구가 필요하다. 둘째, 원인요소 추천 모듈의 경우 원인요소 항목이 149개로 상당히 많아, 인적오류 분석 사례가 적은 경우 유효한 수준 이상의 연관성 값을 도출하기 어려울 것으로 예상된다. 이러한 문제는 인적오류 분석 사례가 축적되어 가면서 어느 정도 해결될 것으로 기대되나 이에 대한 보다 정교한 방법의 개발이 필요하다. 셋째, 유사사고 검색 모듈과 원인요소 추천 모듈에 대해 향후 사고 사례가 일정 개수 이상 모여졌을 때 사용자의 선택사항을 반영하여 좀 더 정교한 추천 및 검색이 가능한 학습 메커니즘에 대한 연구가 필요하다. 마지막으로, 사고 분석을 CAS-HEAR에 적용한 분석 결과와 분석 전문가들과의 인터뷰를 통해 도출된 분석 결과와의 비교 · 분석을 통해 CAS-HEAR의 신뢰성을 검증하는 것이 필요하다.

참고 문헌

Sarwar, B. M., et al., "Analysis of Recommendation Algorithms for E-Commerce", *Proceedings of the 2nd ACM Conference on Electronic Commerce*, pp.158-167, Minneapolis. MN. 2000.

Sarwar, B. M., et al., "Item-based Collaborative Filtering Recommendation Algorithms", *Proceedings of the 10th International World Wide Web Conference*, pp.285-295, Hong Kong. 2001.

Herlocker, J. L., et al., "An Algorithmic Framework for Performing Collaborative Filtering", *Proceedings of the 22nd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp.230-237, Berkeley. CA. 1999.

Reason, J., *Human Error*, New York: Cambridge University Press, 1990.

Resnick, P., et al., "GroupLens: An Open Architecture for Collaborative Filtering of Netnews". *Proceedings of the 1994 ACM Conference on Computer Supported Cooperative Work*, pp.175-186, Chapel Hill. 1994.

KAIST, *안전업무종사자 교육훈련체계 구축 연구보고서*, 1차년도 최종보고서, 2006.

KAIST, *안전업무종사자 교육훈련체계 구축 연구보고서*, 2차년도 최종보고서, 2007.

김동산, 백동현, 윤완철, 인적오류 분석기법의 국내외 활용 현황 및 국내 철도 산업에의 적용 방안, *한국철도학회논문집*, 10(1), pp.7-15, 2007.

김재경, 안도현, 조윤호, 개인별 상품추천시스템, WebCF-PT: 웹마이닝과 상품계층도를 이용한 협업 필터링, *경영정보학연구*, 15(1), pp.63-79, 2005.

신민주, 인적오류 분석을 위한 지원시스템 프레임웍 개발에 관한 연구 – 철도 산업을 중심으로, *한양대학교 석사학위논문*, 2008.

이기현, 고병진, 조근식, 연관규칙 과 협력적 여과 방식을 이용한 추천 시스템, *한국지능정보시스템학회논문지*, 8(2), pp.91-103, 2002.

홍태호, 이희정, 서보밀, 클러스터링 기반 사례기반추론을 이용한 웹 개인화 추천시스템, *한국지능정보시스템학회논문지*, 11(1), pp.107-121, 2005.

◐ 저자 소개 ◑

❖ 신 민 주 ❖ shinmj@syds.com
한양대학교 일반대학원 e-business경영학과 석사
현 재: 삼양 데이터 시스템 e-Biz 팀
관심분야: 데이터 마이닝, AI, 인적오류 분석

❖ 백 동 현 ❖ estarbaek@hanyang.ac.kr
KAIST 산업공학과 박사
현 재: 한양대학교 안산캠퍼스 경상대학 경영학부 부교수
관심분야: 경영정보시스템, Cognitive Systems Engineering

❖ 김 동 산 ❖ kimdongsan@gmail.com
KAIST 산업 및 시스템공학과 석사
현 재: KAIST 산업 및 시스템공학과 박사과정
관심분야: HCI, Cognitive Systems Engineering

❖ 윤 완 철 ❖ wcyoon@kaist.ac.kr
Georgia Institute of Technology 산업 및 시스템공학과 박사
현 재: KAIST 산업 및 시스템공학과, 지적서비스공학과 교수
관심분야: Cognitive Systems Engineering, HRI, HCI

논 문 접 수 일 (Date Received) : 2008년 06월 24일
논 문 수 정 일 (Date Revised) : 2008년 08월 18일
논문게재승인일 (Date Accepted) : 2008년 08월 18일