Journal of the Korea Convergence Society
Vol. 10. No. 8, pp. 15-20, 2019

ISSN 2233-4890
https://doi.org/10.15207/JKCS.2019.10.8.015

텍스트 마이닝 기반의 음악치료 관리 모델에 관한 연구

박성현[1], 김재웅[2*], 김동현[3], 조한진[4]
[1]공주대학교 컴퓨터공학과 박사과정, [2]공주대학교 컴퓨터공학부 교수,
[3]공주대학교 컴퓨터공학과 박사, [4]극동대학교 에너지IT공학과 교수

A Study on the Music Therapy Management Model Based on Text Mining

Seong-Hyun Park[1], Jae-Woong Kim[2*], Dong-Hyun Kim[3], Han-Jin Cho[4]
[1]Ph.D Course, Dept. of Computer Engineering, Kongju National University
[2]Professor, Dept. of Computer Science & Engineering, Kongju National University
[3]Ph.D, Dept. of Computer Engineering, Kongju National University
[4]Professor, Dept. of Energy IT Engineering, Far East University

요 약 음악치료는 장애아동 및 정신치료에 많은 효과를 보이고 있다. 오늘날의 음악치료 시스템은 구체적인 치료 시스템이 구축되어 있지 않은 상황으로, 음악 치료사들이 정확한 치료를 하기 위해 다양한 음악치료 사례들과 치료 이력 데이터들을 분석하고, 해당 환자 또는 내담자에게 가장 적합한 치료를 시행해야 하지만, 현실은 여러 가지 요인들로 인해 많은 어려움이 따르고 있다. 이를 해결하고자 본 논문에서는 기존 치료 데이터와 텍스트 마이닝 기술을 융합한 음악치료 지식관리 모델을 제안한다. 제안 모델을 활용하면 유사한 사례검색 및 환자에 관련된 구체적이고 확실한 데이터들을 기반으로 환자 또는 내담자로 하여금 정확하고 효과적인 치료가 가능하다. 이를 통해서 음악치료의 본래 목적과 그 효과를 최대로 이끌어 내는 효과를 기대할 수 있고 나아가 많은 환자들의 치료에 도움이 될 것으로 기대된다.

주제어 : 융합, 텍스트 마이닝, 음악 치료, 지식관리, 데이터 분석

Abstract Music therapy has shown many benefits in the treatment of disabled children and the mind. Today's music therapy system is a situation where no specific treatment system has been built. In order for the music therapist to make an accurate treatment, various music therapy cases and treatment history data must be analyzed. Although the most appropriate treatment is given to the client or patient, in reality a number of difficulties are followed due to several factors. In this paper, we propose a music therapy knowledge management model which convergence the existing therapy data and text mining technology. By using the proposed model, similar cases can be searched and accurate and effective treatment can be made for the patient or the client based on specific and reliable data related to the patient. This can be expected to bring out the original purpose of the music therapy and its effect to the maximum, and is expected to be useful for treating more patients.

Key Words : Convergence, Text Mining, Music Therapy, Knowledge Management, Data Analysis

*Corresponding Author : Jae-Woong Kim(jykim@kongju.ac.kr)
Received July 2, 2019 Revised August 9, 2019
Accepted August 20, 2019 Published August 28, 2019

1. 서론

음악치료에서의 음악치료사의 역할은 환자의 치료목표에 도달하기 위해 가장 중요시되는 요인이다. 음악치료사에게 필요한 자질은 다양한 음악에 관련된 전문적인 기술, 환자에 대한 지식, 환자의 치료 목적에 맞는 자료수집과 같은 치료에 관련된 기술, 동료 직원과 환자의 보호자들과의 효율적인 교류 및 기본적인 직업윤리와 같은 임상기술로 구분할 수 있다. 특히 음악치료사의 전문성에는 전문 분야 학문에 대한 많은 이해를 기반으로 실질적으로 임상에 적용하는 전 과정이 포함된다[1].

음악치료사들은 개인의 건강 회복이 목표이기에 전문성이 요구되는 직업으로, 음악치료의 대상은 유아에서 노인에 이르기까지 다양하며, 각 대상의 주요한 증상과 필요에 의하여 음악치료의 목표 및 목적이 다르게 적용되어야 한다. 지적장애를 대상으로 일반적으로 활용되는 치료 매체 중 하나인 음악치료는 선율 및 가사를 직접 구성하는 기법을 포함하는 치료적 노래 만들기를 통하여 자아 존중감 향상 및 자기표현 능력을 위한 중재로 활용하고 있는 실정이다[2].

음악치료의 최종 목적인 환자에 대한 치료 목표에 도달하기 위해서는 음악치료사의 완벽한 자질과 능력이 요구된다. 그러나 실질적으로는 사회적인 환경과 다양한 여건들로 인해 완벽할 수 없는 것이 현실로서, 음악치료사들은 내담자에게 필요한 것을 적절하게 제공해야 하는 문제로 실제 임상에서 많은 불안을 경험하고 있다[3]. 특히 상담과 심리치료에서 동기는 치료의 성공여부가 결정되는 핵심적인 요소라 할 수 있고, 잘 동기화된 내담자일 경우 치료의 효과성에 대한 기대가 높다[4].

본 논문에서는 텍스트 마이닝 기술을 음악치료에 적용하여 보다 정확하고 가장 알맞은 음악치료를 내담자에게 제공하는 음악치료 지식관리 모델을 제안한다. 텍스트 마이닝을 음악치료 부분에 적용하게 되면 음악치료에 관련된 데이터들을 분류 및 군집화처리, 원하는 음악치료 데이터 또는 내담자에게 가장 적합한 음악치료 데이터 검색 및 추출이 가능하다.

제안 모델은 내담자에 가장 적합한 음악치료 데이터와 유사 사례들을 검색 및 추출하여 음악치료사에게 제시하는 시스템으로서, 향후 음악치료에 관련된 데이터들을 보다 체계적이고, 구체적으로 데이터를 관리할 수 있으며, 음악치료의 활성화와 음악치료 목적 달성에 많은 도움이 될 것으로 기대된다.

2. 관련연구

2.1 Text Mining

대량의 비정형, 반정형 데이터의 텍스트를 분석하고, 이를 통하여 정보를 추출하고, 텍스트 기반 데이터베이스에 있는 데이터로부터 유용하고, 필요한 정보를 추출하고, 이를 통하여, 지식과 패턴을 발견하여 사용자에게 필요한 정보를 자동 추출하여 제공하는 텍스트 처리 기술과 처리과정이 텍스트 마이닝이다[5].

웹을 기반으로 생성되는 데이터의 양이 나날이 증가하고 있고, 비정형 데이터 분석에 대한 연구가 활발하게 이루어지고 있어, 텍스트 마이닝에 대한 중요성이 나날이 증가되고 있는 상황이다[6]. 텍스트 마이닝은 웹상에서 사용자의 데이터 검색에 도움을 주고, 대량의 데이터에 대한 재해석, 문서의 군집화 및 분류, 내용 요약, 문석의 색인과 추천, 번역, 대표 키워드 추출, 많은 양의 문서 탐색 등이 대표적인 응용분야이다[7]. 또한 텍스트 마이닝은 신문이나 논문, 스팸 처리, 보고서 요약, 질의 응답 시스템, 대표적인 키워드나 토픽 추출, 데이터베이스에서 문서의 군집화 및 분류, 특정 주제 논문 검색, 정보 추출 등 다양한 분야에서 활용되고 있다[8].

특히 주제어를 기반으로 하여 네트워크 분석을 통한 국내 감정노동의 연구 동향 탐색[9], 만성질환자를 대상으로 모바일 헬스 중재 연구 동향 분석[10], 포털사이트에 게시된 인터넷 뉴스를 수집하여 부동산 시장의 동향 분석 연구[11], 대학과 공공연구소에서 최근 개발한 기술을 기반으로 4차 산업혁명 핵심기술에 대한 연관 분석 연구[12]와 같이 많은 분야에서 연구가 이루어지고 있다.

2.2 TF-IDF

TF-IDF(Term Frequency Inverse Document Frequency)는 텍스트 마이닝 및 정보 검색 분야에서 활용되고 있는 가중치로서, 많은 문서들로 이루어진 문서들 중에 특정 단어가 특정 문서 내에서 중요도가 얼마나 높은지를 나타내는 통계 기반의 기술로서[13-15], 검색 결과의 순위 결정, 키워드 추출 및 문서 유사도 측정에 활용되고 있다.

TF는 단어 빈도수를 타나내는 것으로 문서 내에서 단어가 얼마나 자주 등장했는지를 나타내는 값이며, IDF는 DF값의 역수 값으로 문서 집합에서 특정한 단어가 얼마나 자주 등장했는지를 표현한 값으로, TF는 문서 내의 단어들에 대한 총 빈도수를 사용하여 계산할 수 있으나,

문서의 단어수가 많거나 클 경우 지속적으로 값이 커지는 문제를 내포하고 있다[16].

다음의 (식 1)은 TF-IDF를 수식으로 표현한 것으로 d는 문서, t는 단어이고, $tf(t,d)$는 문서 d 내에서 단어 t의 총 빈도수이며, 다음의 (식 2)는 IDF를 나타낸 것으로 D는 전체 문서로서 단어 t가 포함되어 있는 문서의 수로 나누고, log를 통하여 값을 획득할 수 있다[17].

IDF에 log를 취하는 것은 IDF의 값이 무한대로 커지는 것을 방지하기 위하여 값을 줄이는 것이다.

$$tfidf(t,d,D) = tf(t,d) \times idf(t,D) \qquad (식\ 1)$$

$$idf(t,D) = \log \frac{|D|}{|d \in D : t \in d|} \qquad (식\ 2)$$

3. 제안 모델

3.1 모델 구성도

다음의 Fig. 1은 제안 모델의 구성도이다.

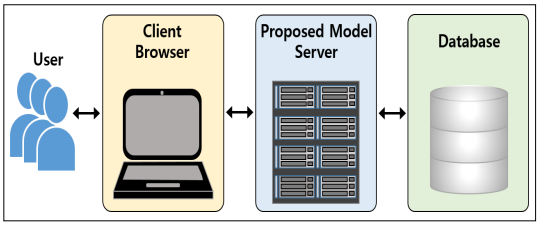

Fig. 1. Proposed Model Configuration

음악치료 지식관리 모델의 서버에서는 환자 또는 내담자들의 데이터들을 관리하며, 회원가입, 음악 치료 기록 데이터 업로드, 검색 등과 관련된 작업을 수행한다.

사용자는 의료진들과 음악치료사들로서 최초에 회원가입을 진행하고 필요한 정보의 검색 및 치료 결과 보고서를 업로드를 통하여 사용자들 간의 정보 공유가 가능하며, 필요한 자료를 다운로드하여 담당 환자 또는 내담자에게 가장 적합한 음악치료 데이터를 제공하는 구조로 구성하였다.

3.2 시스템 프로세스

다음의 Fig. 2는 제안 모델의 프로세스로서 크게 4단계로 이루어져 있다.

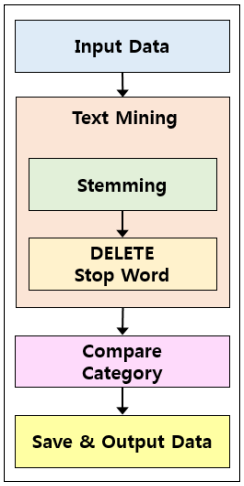

Fig. 2. Model Process

첫째, 데이터 입력 단계로서, 음악치료를 시행할 환자 또는 내담자의 상담 자료를 입력한다.

둘째, 텍스트 마이닝 단계로서, 사용자가 입력한 환자 또는 내담자의 데이터를 기준으로, 텍스트 마이닝을 수행하여 환자 데이터들의 분류 기준이 되는 의미 있는 단어들을 추출한다. 텍스트 마이닝 단계를 수행하기 위해서는 데이터를 분류하기 위한 기준이 되는 단어 사전을 미리 데이터베이스에 생성하고 저장하는 전처리 과정이 필요하다. 단어 사전은 장애아동, 음악치료, 정신장애 음악치료, 노인 음악치료, 의학과 음악과 같이 4가지로 카테고리를 분류하였다.

전처리 과정이 완료된 후 텍스트 마이닝의 첫 번째 단계인 형태소 분석을 수행한다. 입력된 환자 또는 내담자의 데이터에서 어절들의 품사를 파악한 후 명사, 형용사, 동사와 같은 의미있는 품사들을 분리하고, 각각의 단어들에 대하여 단어별로 등장 빈도수를 체크한다.

형태소 분석 과정이 완료된 후, 두 번째로 불용어 제거 과정을 수행한다. 불용어 제거 과정에서는 입력된 데이터에 품사를 부착하고 단위 분할이 완료된 데이터의 분류 기준을 명확히 하기 위하여 불필요한 품사 요소들을 필터링 하여 재추출을 진행한다. 입력된 환자 또는 내담자의 데이터를 토큰화(Tokenization)하여 품사사전에 등록된 품사를 비교 및 검색하여 불필요한 품사들을 제외한 결과를 출력한다. 데이터의 형태소 분석과 불용어 제거 단계의 수행이 완료된 후, 특정한 단어들이 얼마나 자주 등장하는지 확인한다. 이때 등장하는 빈도수가 높을수록 해당 문서에서 중요한 단어로 판단할 수 있다.

다음의 Fig. 3은 TF-IDF를 이용하여 단어별 중요도를 체크하기 위한 알고리즘의 일부분이다.

```
FOR (i ← 0 to P_Data.length step 0)
  NLP = Excute NaturalLanguageProcessing(P_Data);
  Tokens = (NLP)_Excute tokenize;
  FOR (i ← 0 to Token.len step 0)
      IF (no token in P_Data)
          Token_Value + 1;
          Result_Data_Map ADD(Token, Token_Value);
      ELSE (has token in P_Data)
        Get the same token in the Result_Data_Map;
        Token_Value + 1
      END IF
  END FOR
  List ADD Result_Data_Map;
END FOR
```

Fig. 3. TF Algorithm

셋째, 텍스트 마이닝 처리된 데이터의 결과를 전처리 과정에서 등록한 사전에 등록되어 있는 카테고리 데이터와 비교하여 카테고리를 선정하고 분류하는 단계이다.

넷째, 텍스트 마이닝을 통해 분석된 데이터를 정해진 카테고리에 적합한 신규 데이터로 데이터베이스에 저장한다. 저장과 동시에 사용자에게 기존의 환자 또는 내담자 데이터와 입력한 환자 또는 내담자의 데이터를 비교하여 가장 비슷하고 적합한 환자 또는 내담자의 데이터를 검색 및 출력하여 사용자에게 데이터를 제공한다.

3.3 데이터베이스 테이블

다음의 Table 1은 환자 또는 내담자들의 데이터 분류 기준이 되는 공통 카테고리의 정보가 저장된 데이터베이스 테이블의 구조이다.

Table 1. Category DB Table Information

Field	Description	Type
category_code(PK)	category code	INT(11)
category_title	category name	VARCHAR(250)
keywords	category keyword	VARCHAR(500)
create_date	create date	DATETIME
update_date	update date	DATETIME
delete_date	delete date	DATETIME
delete_yn	delete (Y / N)	VARCHAR(1)

텍스트 마이닝 단계를 통하여 해당하는 환자의 데이터에 대한 카테고리 분류가 확정되고, 공통 카테고리 테이블 항목을 기반으로 입력되는 환자의 데이터들을 분류한다.

키워드 필드는 유의어 사전개념으로 분석된 환자 데이터의 단어들이 카테고리와 정확하게 일치하지 않더라도, 키워드에 해당하는 단어들이 각 카테고리에 포함되어 있으면 해당 카테고리로 지정되고, 환자테이블의 카테고리 필드 값이 저장된다.

저장된 필드 값은 해당 환자데이터의 구분자가 되며 추후에 비슷한 환자의 데이터가 사용자로부터 입력 및 검색되었을 경우 사용자에게 같은 카테고리의 해당하는 환자들의 치료이력 정보를 제공한다.

3.4 주요 관리 화면

다음의 Fig. 4는 제안하는 모델의 메인 화면이다. 다양한 관리기능을 이용할 수 있도록 연결 및 제어를 제공하는 포털 시스템 역할을 한다.

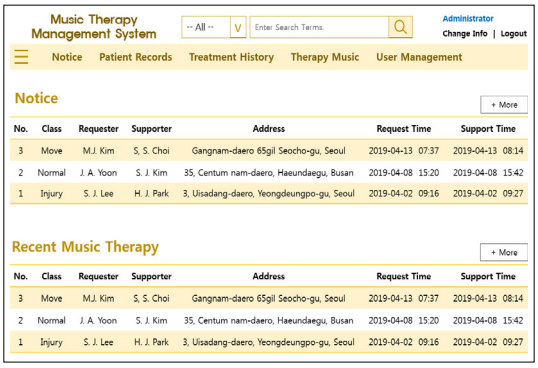

Fig. 4. Proposed Model Main Screen

화면 상단의 각 메뉴별 하위메뉴를 통해서 필요한 기능을 수행할 수 있다. 음악치료 지식관리 모델의 메인화면에서 통합검색 기능을 수행이 가능하며, 다른 환자들의 치료이력을 검색하여 해당 환자의 치료에 도움이 되는 참고 자료를 이용할 수 있다.

다음의 Fig. 5는 환자 및 내담자의 데이터를 입력하는 화면이다. 입력된 환자 및 내담자의 데이터는 텍스트 마이닝 단계를 통하여 데이터의 카테고리가 지정되고 데이터베이스에 저장된다.

Fig. 5. Data Entry Screen

4. 분석 및 고찰

본 논문에서 제안한 모델을 30일간의 테스트 기간 동안 실제 운영하여 사용하면서, 사용자들의 만족도 및 시스템의 효과, 효율에 대한 설문조사를 진행하여 제안 모델의 서비스 평가를 하였다. 다음의 Table 2는 서비스 평가 내용을 정리한 것이다.

Table 2. Service Evaluation Contents

Division	Contents
Date	2019. 01. 12 ~ 2019. 02. 11
Participant's AGE	20 ~ 50
Participant	Medical staff & Music therapist 50 people

음악치료 지식관리 모델 사용에 대한 평가를 위하여 만족도, 효과성, 효율성으로 항목을 분류하였고, 각 항목에 대한 세부 평가 문항을 작성하여 협회 및 병원에 종사하는 음악치료사 및 의료진 40명을 대상으로 설문조사를 수행하였다. 다음의 Table 3은 설문조사 결과를 정리한 것이다.

Table 3. Survey Results

Division	Satisfaction	Efficiency	Effectiveness
Very Good	28	31	29
Good	7	6	7
Normal	3	2	2
Dissatisfaction	2	1	2
Very Dissatisfied	0	0	0

음악치료 지식관리 모델의 설문조사 결과 시스템의 품질이나, 서비스의 효과 및 효용, 사용자 만족도는 만족이상이 평균 89.17%로 제안 모델의 시스템에 대한여 긍정적인 것으로 나타났으며, 보완 사항으로는 기존의 치료이력 데이터의 수정과 보완 부분에 있어 사용자 입력 용이성에 대한 의견이 도출되었다.

5. 결론 및 향후 연구 방향

음악치료는 개인의 치료 목적을 달성하기 위하여 음악적 경험을 기반으로 하는 체계적이고 과학적인 과정이다. 음악, 고객, 음악치료사는 서로 다양한 관련성을 가지고 발전하고 있고, 효과적인 치료 목표를 달성하기 위하여 노력하고 있다. 그러나 음악치료와 관련된 종사자들은 지식 관리 시스템의 부재로 인하여 치료 정보의 공유가 매우 어려운 상황이다. 본 논문에서는 음악 치료와 관련된 치료 방법과 같은 정보를 공유하기 위한 음악치료 지식관리 모델을 제안한다. 이를 통하여 음악치료사가 환자를 치료함에 있어, 효과적인 치료가 이루어질 것으로 기대된다. 향후 연구방향으로는 음악치료사들 및 의료진들의 보완 사항인 치료이력 데이터의 수정 및 활용 부분의 기능을 용이하게 하고, 비전문가들도 음악치료의 정보를 공유하고 보급될 수 있는 연구가 계속되어야 할 것이다.

REFERENCES

[1] H. J. Chong. (2007). Intrapersonal competency in music therapy: Challenges for education and clinical training. *Journal of Music and Human Behavior, 4(2),* 106-114.

[2] Y. E. Yun. (2018). Improved self-expression of adolescents with cochelar implants after a songwriting intervention: A case study. *Journal of Music and Human Behavior, 15*(2), 69-92.
DOI : 101.21187/jmhb.2018.15.2.069

[3] B. L Wheeler. (2002). Experiences and concerns of students during music therapy practica. *Journal of Music Therapy, 39(4), 274-304.*

[4] K. S. Kang. (2012). A Study on the Grounded Theory of Motivation in Clinical Music Therapy Practice for the Mental Rehabilitation Members. *Korean Journal of Music Therapy Education, 9(1),* 37-53.
DOI : 10.21187/jmhb.2012.9.1.037

[5] H. N. Chung, D. H. Kim & B. O. Goh. (2004). The Development and Application Of Cyber Counseling

System for the Gifted Class. *Journal of the Korean Association of Information Education, 8(2),* 177-187.

[6] S. G. Cho & S. B. Kim. (2012). Finding Meaningful Pattern of Key Words in IIE Transactions Using Text Mining. *Journal of the Korean Institute of Industrial Engineers, 38(1),* 67-73.
DOI : 10.7232/JKIIE.2012.38.1.067

[7] Y. K. Jeong. (2016). *An exploratory study of consumer counseling data using text mining : focusing on cancer insurance.* Master's Thesis. Sookmyung Women's University, Seoul.

[8] J. Wang. J. Liu, & C. Wang. (2007). Keyword extraction based on PageRank. *In Pacific-Asia Conference on Knowledge Discovery and Data Mining.* (pp. 857-864). Berlin, Heidelberg : Springer.

[9] N. Y. Lee, J. H. Kim, H. J. Mun. (2019). Exploration of Emotional Labor Research Trends in Korea through Keyword Network Analysis. *Journal of Convergence for Information Technology, 9(3),* 68-74.
DOI : 10.22156/CS4SMB.2019.9.3.068

[10] Y. J. Son, S. K. Lee. (2019). Analysis of research trends on mobile health intervention for Korean patients with chronic disease using text mining. *Journal of Digital Convergence, 17(4),* 211-217.
DOI : 10.14400/JDC.2019.17.4.211

[11] H. J. Chun. (2019). Analysis of Real Estate Market Trend Using Text Mining and Big Data. *Journal of Digital Convergence, 17(4),* 49-55.
DOI : 10.14400/JDC.2019.17.4.049

[12] H. Ryu & Y. Y. You. (2018). The Fourth Industrial Revolution Core Technology Association Analysis Using Text Mining. *Journal of Digital Convergence, 16(8),* 129-136.
DOI : 10.14400/JDC.2018.16.8.129

[13] D. S. Park & H. J. Kim. (2018). A Proposal of Join Vector for Semantic Factor Reflection in TF-IDF Based Keyword Extraction. *The Journal of Korean Institute of Information Technology, 16(2),* 1-16.
DOI : 10.14801/jkiit.2018.16.2.1

[14] E. S. You, G. H. Choi & S. H. Kim. (2015). Study on Extraction of Keywords Using TF-IDF and Text Structure of Novels. *Journal of the Korea Society of Computer and Information, 20(2),* 121-129.
DOI : 10.9708/jksci.2015.20.2.121

[15] W. S. Choi & S. B. Kim. (2015). N-gram Feature Selection for Text Classification Based on Symmetrical Conditional Probability and TF-IDF. *Journal of the Korean Institute of Industrial Engineers, 41(4),* 381-388.
DOI : 10.7232/JKIIE.2015.41.4.381

[16] D. S. Park & H. J. Kim. (2016). A Proposal of Join Vector for Semantic Factor Reflection in TF-IDF Based Keyword Extraction. *The Journal of Korean Institute of Information Technology, 16(2),* 1-16.
DOI : 10.14801/jkiit.2018.16.2.1

[17] J. H. Kim. (2019). *Method of Keyword Recommendation Considering Importance and Correlation of Words.* Master's Thesis. Chosun University, Gwangju.

박 성 현(Seong-Hyun Park) [정회원]

· 2012년 2월 : 충남대학교
· 2017년 2월 : 공주대학교 IT공학전공(공학석사)
· 2019년 7월 ~ 현재 : 공주대학교 컴퓨터공학과 박사과정
· 관심분야 : 인공지능, 컴퓨터음악, 빅데이터
· E-Mail : giornopark@daum.net

김 재 웅(Jae-Woong Kim) [종신회원]

· 1983년 2월 : 중앙대학교 전자계산학과(공학사)
· 1988년 2월 : 중앙대학교 대학원 컴퓨터공학과(공학석사)
· 2000년 2월 : 대전대학교 대학원 컴퓨터공학과(공학박사)
· 1992년 8월 ~ 현재 : 공주대학교 컴퓨터공학부 교수
· 관심분야 : 소프트웨어공학, 빅데이터, 멀티미디어공학
· E-Mail : jykim@kongju.ac.kr

김 동 현(Dong-Hyun Kim) [정회원]

· 1986년 2월 : 중앙대학교 전기공학과(공학사)
· 2005년 2월 : 공주대학교 컴퓨터멀티미디어공학과(공학석사)
· 2010년 2월 : 공주대학교 컴퓨터공학과(공학박사)
· 관심분야 : 인공지능, 영상처리, 지식관리, 시뮬레이션
· E-Mail : dhkim977@naver.com

조 한 진(Han-Jin Cho) [종신회원]

· 1999년 2월 : 한남대학교 컴퓨터공학과(공학석사)
· 2002년 8월 : 한남대학교 컴퓨터공학과(공학박사)
· 2002년 8월 ~ 현재 : 극동대학교 에너지IT공학과 교수
· 관심분야 : 융복합 보안, 모바일 보안, 네트워크 보안
· E-Mail : hanjincho@hotmail.com