

빅데이터 처리 기반의 범죄 예방 스마트 시스템에 관한 연구

김원
우송대학교 IT융합학부 교수

A Study on the Crime Prevention Smart System Based on Big Data Processing

Won Kim
Professor, Division of IT Convergence, Woosong University

요약 4차산업혁명 이후, 빅 데이터 분석, 로봇공학, 사물인터넷, 인공지능 등의 핵심 기술들이 여러 분야에서 활용되고 있다. 일반적으로 빅 데이터 기술은 방대한 데이터를 모으고, 분석하고 처리하며, 공급하는 단계를 갖는 것으로 이해되고 있다. 유용한 방대한 데이터 중에 하나인 범죄 기록은 현재까지는 범죄가 일어난 후 수사 정보를 획득하는데 이용이 되고 있다. 방대한 범죄 기록을 빅 데이터 처리하여 범죄를 예측하는 데 사용할 경우 범죄의 발생 빈도를 줄일 수 있을 것으로 판단된다. 본 논문에서는 범죄 기록을 빅 데이터 처리하여 범죄 발생 확률을 스마트 기기 사용자에게 제공하는 스마트 시스템을 제안한다. 구체적으로 기기상의 전자지도에 범죄 발생 확률을 나타내어 안전한 이동 경로를 안내하는 시스템을 의미한다. 소규모 지역을 다루는 스마트 앱으로 실험한 결과 범죄 예방에 있어서 활용성이 비교적 좋은 것으로 판단된다.

주제어 : 융합, 빅 데이터, 범죄 예방, 데이터 처리, 위치 기반 접근

Abstract Since the Fourth Industrial Revolution, important technologies such as big data analysis, robotics, Internet of Things, and the artificial intelligence have been used in various fields. Generally speaking it is understood that the big-data technology consists of gathering stage for enormous data, analyzing and processing stage and distributing stage. Until now crime records which is one of useful big-sized data are utilized to obtain investigation information after occurring crimes. If crime records are utilized to predict crimes it is believed that crime occurring frequency can be lowered by processing big-sized crime records in big-data framework. In this research the design is proposed that the smart system can provide the users of smart devices crime occurrence probability by processing crime records in big-data analysis. Specifically it is meant that the proposed system will guide safer routes by displaying crime occurrence probabilities on the digital map in a smart device. In the experiment result for a smart application dealing with small local area it is showed that its usefulness is quite good in crime prevention.

Key Words : Convergence, Big Data, Crime Prevention, Data Processing, Location based Approach

*This research was supported by the MISP(Ministry of Science, ICT), Korea, under the National Program for Excellence in SW(2018-0-01879) supervised by the IITP(Institute of Information & communications Technology Planning & Evaluation)(2018-0-01879).

*Corresponding Author : Won Kim(kimwon@wsu.ac.kr)

Received October 6, 2020

Accepted November 20, 2020

Revised November 5, 2020

Published November 28, 2020

1. 서론

4차산업혁명을 통하여 산업 전반의 발전과 더불어 컴퓨터를 기반으로 하는 산업 간의 융합이 급속하게 이루어지고 있고, 데이터 분석, 로봇공학, 사물인터넷, 인공지능 등의 핵심 기술들이 다시 조망을 받으며 여러 분야에서 활용되고 있다. 그 중 빅 데이터 분석 기술은 단말기 접속 서버의 발전, 휴대기기의 확대 보급 등으로 인하여 데이터를 획득하는 작업이 시스템적으로 가능해지면서 각광을 받고 있다. 최근 보이스 피싱과 관련된 사례를 빅 데이터로 처리해서 공격기법을 분석하고 대응방안을 찾는 시도가 있었고[1], 대량의 금융거래 기록 분석에 적용하여 비정상적인 이상금융거래를 탐지하는 연구가 있었다[2]. 또한 기상 분석 서버를 운영할 때 발생하는 대량의 기상 데이터를 분석하여 재해에 대한 적절한 대비 전략을 수립하는 데 이용되기도 했다[3]. 한편 의학 분야에서는 컨볼루션 기반의 인코더를 디블링 프레임워크로 기계 학습하여 폐종양을 진단하려는 시도가 있었고[4], 방대한 줄기세포 영상을 빅 데이터 분석하여 유사분열의 자동 탐지 기법의 연구를 시도하였다[5]. 이를 더 발전시켜 인간의 두뇌와 컴퓨터를 인터페이스하는 시도로서 방대한 뇌파도 신호를 컨볼루션 기반의 신경회로망으로 학습시키는 연구가 있었다[6]. 통신분야에서는, 통신채널의 다중경로 설계에 활용되기도 했으며[7-9], 싱가포르에서는 수백만 건의 휴대폰 통화 기록을 분석 처리하여 대단위 교통망 설계에 이용하는 연구를 시도하였다[10]. 연구에서 시도되는 빅 데이터 기법은 태생적 특성으로 비교적 많은 처리 용량을 필요로 하는데, 맵리듀스 기반의 처리 기법이 고안되어 저장 효율성을 높이려는 연구가 있었다[11]. 범죄 수사 분야에서도 이러한 기술들을 기반으로 다양한 연구가 이루어지고 있으며, 수만 건의 범죄 기록을 분석하여 범죄 수사 단서를 찾는 데 이용이 되었고[12], 비교적 큰 단위의 통화 기록을 분석 처리하여 범죄

수사에 이용한 연구가 있었다[13]. 그러나 범죄 분야에서 범죄 데이터를 수사에 이용하는 것도 의미가 있으나, 범죄 예측을 하여 그 발생빈도를 저감시키는 것이 사회 안전에 의미가 더 크다고 할 수 있으며, 이를 위해 빅 데이터 분석 기술과 데이터 마이닝을 결합하여 범죄 동향을 필요한 조건에 따라 예측하려는 연구가 진행되었다[14,15]. 본 논문에서는 과거의 범죄 기록을 대상으로 데이터베이스를 구축하고, 빅데이터 분석 처리를 기반으로 국민들의 시간 정보, 위치정보, 연령, 성별과 같은 필수 정보를 결합하여 현재 위치에서 범죄 위험도를 스마트 기기를 통하여 고지할 수 있는 범죄 예방 스마트 시스템을 제안한다.

2. 연구 동기

발생한 범죄에 대한 수사를 정확하게 하는 것이 중요하지만, 한발 나아가 예측하여 발생 빈도를 낮추는 것이 사회 구성원의 행복지수 유지와 향상에 필요하다. 다음의 Fig. 1과 같이 같이 범죄 기록(Crime record)은 문서 또는 전자문서 형태로 방대한 분량으로 수사 기관에 보관되어 있다. 이 기록은 일반적으로 범죄가 발생하면 수사에 적극적으로 활용되지만, 사전에 특성을 분석하여 예방에 이용하기에는 그 크기의 방대함으로 인해 기술적인 어려움이 있다. 범죄 기록을 모아 데이터베이스로 구축하여 빅데이터 분석 처리를 하고, 현재의 시간정보, 위치정보, 연령, 성별 등 소프트웨어 사용자의 필수정보와 결합하면 스마트 기기의 전자 지도에서 현 위치에 대한 범죄 확률을 산출하여 사용자에게 제공할 수가 있다. 빅데이터 분석 기술과 범죄 데이터베이스를 결합하고 이를 스마트 기기의 앱을 통해 사용자에게 현재 위치의 범죄 확률을 제공함으로써, 높은 범죄가 발생한 위치를 우회하여 다닐 수 있게 하는 방법으로 범죄 발생 빈도를 낮추려

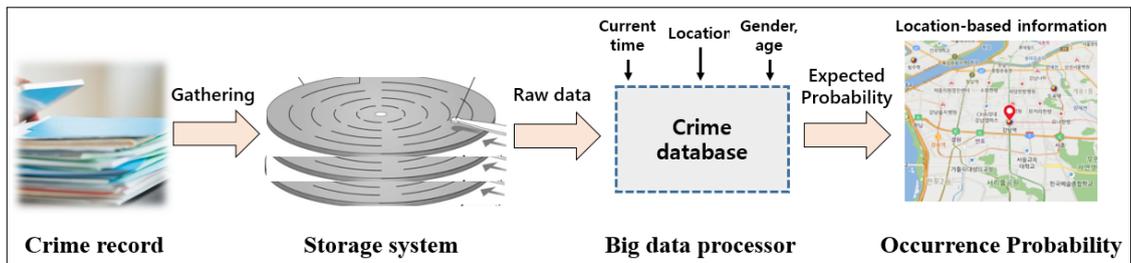


Fig. 1. Concept of big data processing for crime prevention

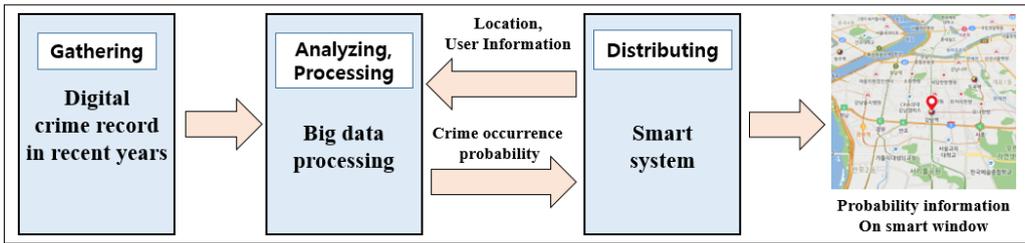


Fig. 2. Overall configuration for the proposed system

는 의도가 이 연구의 주된 동기이다.

3. 범죄 예방 스마트 시스템

빅데이터 분석처리 기술은 데이터를 모으는 것 (Gathering), 분석(Analyzing)하고 처리(Processing)하는 것, 공급(Distributing)하는 것 등의 크게 네 가지 기술요소를 가지고 있다. 앞의 Fig. 2는 제안 시스템의 전체적인 구성도로서, 빅 데이터 기술 구성요소와 스마트 기기와의 연결성을 고려하여 설계한 것이다. 지난 수 년 간의 범죄 기록을 확보하여 이를 데이터베이스로 구축하고, 분석 및 처리하여 확률적 데이터로 변환한다. 확률적 데이터의 단순한 모음은 활용의 측면에서 무의미하기 때문에 사용자 현재 위치와 같은 적절한 정보와 연관하여 범죄 발생 확률을 계산하여 스마트 시스템에 공급하는 개념이 전반적인 데이터 흐름이다.

전체 시스템 구성 개념을 세부적으로 제안하는 시스템의 설계에 적용하면 다음의 Fig. 3과 같이 나타낼 수가 있다. 단말기를 통해 범죄 기록을 적절하게 가공하고, 이를 편집하여 지난 수 년 간의 방대한 범죄 데이터베이스를 구축한다. 여기까지가 데이터를 모으는 단계 (Gathering stage)이다. 이 범죄 정보(Crime data)를 빅데이터 분석 기법으로 처리하여 범죄에 대한 발생 확

률 정보를 정리한다(Analyzing and processing stages). 스마트 기기의 앱을 이용하여 사용자의 현재 위치, 연령, 성별 정보 등을 담아 빅데이터 처리 서버에 접속하면 현재 서버에 접속하면 현재 위치를 기반으로 범죄 발생 빈도가 높은 지점들을 네트워크로 받아 전자지도 위에 표시가 가능하다(Distributing stage). 또한 여성 범죄 또는 고령층 범죄 등 세부적으로 범죄를 구분하여 범죄 발생 확률을 표시할 수 있다. 스마트 앱을 사용하는 사람은 자신의 현재 위치를 기준으로 주변에 어떤 종류의 범죄가 어떤 확률로 발생했는지 판단하여 본인이 원하는 목적지까지 안전하게 도달할 수 있는 이동 경로를 제공받을 수 있다.

예를 들어 어떤 사용자가 잘 알지 못하는 지역으로 출장을 갔을 경우, 숙소를 찾아가기 위해 임의의 길과 방향을 선택하기 보다는 범죄 예방 앱을 사용하여 빅 데이터 처리 결과 측면에서 비교적 안전한 경로를 선택하는 것이 나쁜 일을 당할 확률을 낮출 수 있기 때문이다. 스마트 폰과 같은 지능형 단말 장치에서 전자지도에 빅 데이터 처리 결과를 효과적으로 나타내기 위해서는 앱 구조 설계가 필요하다. 다음의 Fig. 4는 제안 시스템의 앱 소프트웨어 구조이다. 커널 수준, 라이브러리 수준, 프레임워크 수준, 어플리케이션 수준 등 크게 4가지 수준 (Level)로 나누어 앱을 설계하였다. 커널 수준의 와이파이 드라이버(Wi-fi driver)는 빅 데이터 처리 서버와 네

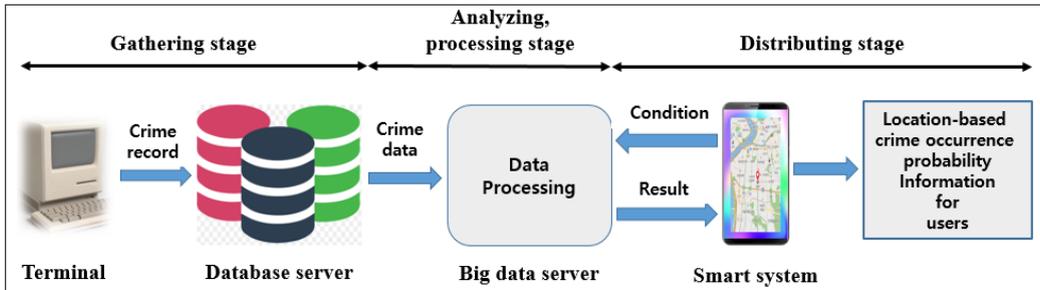


Fig. 3. System design of the proposed system

트위크로 연결되는 지점으로, 이를 통해 사용자 정보가 요구 패킷(Request packet)에 담겨 빅 데이터 처리 서버로 전송되고, 처리 결과로서 범죄 확률 정보가 수신된다. 범죄 확률 정보는 애플리케이션 수준의 앱 소프트웨어에서 처리하여 그래픽 라이브러리의 지원을 받아 그래픽 지도 정보를 구성하게 되고, 최종적으로 그래픽 메모리로 전송되어 디스플레이 드라이버에서 지도정보 형태로 사용자가 볼 수 있게 된다.

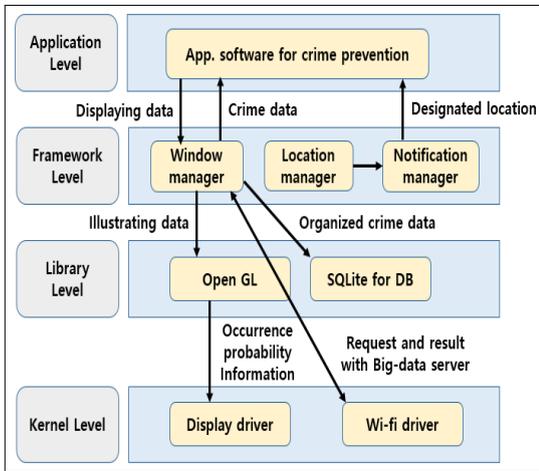


Fig. 4. Application software structure design of the proposed system

다음의 Table 1은 제안하는 시스템의 설계 요구사항을 정리한 것이다.

Table 1. Design requirements for the proposed system

Configuration	Required specification
Big data analytics	AppEngine, Hadoop on Azure
Relational database	Cloud SQL, SQL Azure, Oracle
Streaming processing	Search API, Table storage
Developing Language	Java, Kotlin, C++
Developing app. tool	Eclipse, Android studio
Supporting device	Android phone, smart watch
Android system	Above Android 10.0(Q)

기계학습 알고리즘 중 하나인 의사결정나무모형을 선정하여 예측 모델을 구축 하였으며, 이를 기반으로 주요 5대 범죄를 대상으로 시나리오를 작성하여 범죄 발생 위험지역을 예측하는 빅데이터 분석 기법을 적용하였다. 빅

데이터 분석에는 AppEngine, Hadoop 등의 엔진을 사용 하였으며, 관계형 데이터베이스는 Cloud SQL이나 Azure를 활용하였고, 안드로이드 운영체제는 최소한 Android 10.0 이상 수준인 것으로 고려하여 설계하였다.

4. 스마트 앱 구현 및 평가

4.1 제안 시스템 구현

실험을 위해 제안된 시스템의 설계 내용을 고려하여 스마트 앱을 구현하였으며, DJ시 D구의 특정한 동에 대해 3년간의 가상 범죄 데이터를 구축하고, 이를 처리하여 범죄 발생 확률 정보를 구성하였다. 가상데이터는 다음의 Table 2와 같이 관찰정보 필드, 발생시간 필드, 위치정보 필드, 범죄종류 필드 등으로 범죄의 종류와 상태를 파악할 수 있도록 구성되어 있다. 범죄종류 필드의 5대 범죄는 살인, 강도, 성범죄, 절도, 폭력을 의미한다. 범죄 발생 지역은 지도상에 시각적 표현이 가능하도록 전체주소에 시군구와 읍면동 등으로 구분하여 데이터 셋을 구성하였으며, 범죄 발생 지역의 담당 경찰서와 지구대 정보를 포함하였다.

Table 2. Configuration of Crime Data Set

Crime control field		Crime occurrence time field	
Police station name	Patrol division name	Occurrence date	Occurrence time
Crime location filed			Classification field
Occurrence city	Detailed location	Entire address	5 major crime

다음의 Fig 5는 제안 시스템의 동작 화면으로서, 현재 사용자 위치를 기준으로 주변의 범죄 발생 확률을 시각적으로 확인할 수 있고, 범죄 발생 확률이 낮은 이동경로 정보를 제공받아 실시간으로 확인할 수 있다.

4.2 제안 시스템 평가

제안 시스템인 범죄 예방 스마트 앱 시스템의 평가를 위하여 시스템 개발자, 경찰 관계자, 일반 사용자 50명을 대상으로 설문조사를 실시하였다. 다음의 Table 3은 설문 조사결과를 분석하여 정리한 것이다.



Fig. 5. Design result of user interface

Table 3. Survey Results

Classification	User satisfaction	effectiveness	efficiency
Very good	31	29	35
Good	12	15	11
Moderate	4	4	3
Dissatisfied	2	2	1
Very Dissatisfied	1	0	0

제안 시스템을 30일간 사용하게 하여, 설문 조사 대상자들의 만족도를 평가하였다. 설문 대상자들의 연령은 10대에서 60대까지 스마트 기기를 사용하고 있는 사용자들을 대상으로 하였다. 평가 결과를 분석한 결과 '좋다' 이상의 의견이 88.66%로 제안 시스템에 대한 사용자 만족은 전반적으로 높게 평가되었으며, 앱을 처음 사용하는 경우 익숙해지는데 시간이 필요하다는 의견과 범죄 통계 및 분석에 대한 기능 추가에 대한 의견이 도출되었다.

5. 결론 및 향후 연구 방향

현대 사회가 발전하면서 사람들은 편리한 생활을 하고 있지만, 반면에 범죄의 발생은 끊임없이 우리 주변에서 일어나고 있다. 빅 데이터 처리 기술은 범죄 발생 이후에 정확한 수사에 활용되고 있으며, 현재까지 일어난 방대한 분량의 범죄 기록을 분석, 처리하여 위치 기반의 범죄 발생 확률 정보의 획득이 가능하다. 현대인이 누구나 가지고 있는 스마트 기기의 전자지도 위에 표현하여 사용자의 안전을 적극적으로 도모할 수 있다. 본 논문은 빅 데이터 처리 기법을 응용한 범죄 예방 스마트 시스템을 설계하였고, 구체적으로는 소프트웨어 구조 설계와 설계 요구사항 등을 도출하였다. 소규모 지역에 대한 스마트 앱 실험을 통해 연구의 효용성을 확인하였다. 향후 범죄 기록에 대한 다양한 정보 획득과, 위치 정보 및 사생활 보호에 관한 연구가 계속되어야 할 것이다.

REFERENCES

- [1] J. Kim, J. Go & K. Lee. (2015). A Scheme of Social Engineering Attacks and Countermeasures Using Big Data based Conversion Voice Phishing. *Journal of the Korea Convergence Society*, 6(1), 85-91. DOI : 10.15207/JKCS.2015.6.1.085
- [2] J. Kang, J. Lee & Y. You. (2017). A Study on Implementation of Fraud Detection System (FDS) Applying BigData Platform. *Journal of the Korea Convergence Society*, 8(4), 19-24. DOI : 10.15207/JKCS.2017.8.4.019
- [3] B. Shin & H. Jeon. (2018). A Study on Disaster Information Support using Big Data. *Journal of the Korea Convergence Society*, 9(8), 25-32. DOI : 10.15207/JKCS.2018.9.8.025
- [4] M. Chen, X. Shi, Y. Zhang, D. Wu & M. Guizani. (2017). Deep Features Learning for Medical Image Analysis with Convolutional Autoencoder Neural Network. *IEEE Transactions on Big Data*, 1-1. DOI : 10.1109/TBDDATA.2017.2717439
- [5] A. Liu, Y. Lu, M. Chen & Y. Su. (2017). Mitosis Detection in Phase Contrast Microscopy Image Sequences of Stem Cell Populations: A Critical Review. *IEEE Transactions on Big Data*, 3(4), 443-457. DOI : 10.1109/TBDDATA.2017.2721438
- [6] M. Hosseini, D. Pompili, K. Elisevich & H. Soltanian-Zadeh. (2017). Optimized Deep Learning for EEG Big Data and Seizure Prediction BCI via Internet of Things. *IEEE Transactions on Big Data*, 3(4), 392-404. DOI : 10.1109/TBDDATA.2017.2769670

- [7] Y. Li, J. Zhang, Z. Ma & Y. Zhang. (2020). Clustering Analysis in the Wireless Propagation Channel with a Variational Gaussian Mixture Model. *IEEE Transactions on Big Data*, 6(2), 223-232.
DOI : 10.1109/TBDDATA.2018.2840696
- [8] J. Huang, C. Wang, L. Bai, J. Sun, Y. Yang, J. Li, O. Tirkkonen & M. Zhou. (2020). A Big Data Enabled Channel Model for 5G Wireless Communication Systems. *IEEE Transactions on Big Data*, 6(2), 211 - 222.
DOI : 10.1109/TBDDATA.2018.2884489
- [9] E. P. Xing, Q. Ho, W. Dai, J. K. Kim, J. Wei, S. Lee, X. Zheng, P. Xie, A. Kumar. (2015). Petuum: A New Platform for Distributed Machine Learning on Big Data. *IEEE Transactions on Big Data*, 1(2), 49-67.
DOI : 10.1109/TBDDATA.2015.2472014
- [10] S. Jiang, J. Ferreira & M. C. Gonzalez. (2017). Activity-Based Human Mobility Patterns Inferred from Mobile Phone Data: A Case Study of Singapore. *IEEE Transactions on Big Data*, 3(2), 208-219.
DOI: 10.1109/TBDDATA.2016.2631141
- [11] H. Lee, Y. Kim, & K. Kim. (2018). Implement of MapReduce-based Big Data Processing Scheme for Reducing Big Data Processing Delay Time and Store Data. *Journal of the Korea Convergence Society*, 9(10), 13 - 19.
DOI : 10.15207/JKCS.2018.9.10.013
- [12] J. Hu. (2018). Big Data Analysis of Criminal Investigations. *2018 5th International Conference on Systems and Informatics (ICSAI)*.
DOI : 10.1109/ICSAI.2018.8599305
- [13] S. Khan, F. Ansari, H. A. Dhalvelkar & S. Computer. (2017). Criminal investigation using Call Data Records (CDR) through Big Data technology. *2017 International Conference on Nascent Technologies in Engineering (ICNTE)*.
DOI: 10.1109/ICNTE.2017.7947942
- [14] S. Sathyadevan, M. S. Devan, S. S. Gangadharan. (2014). Crime analysis and prediction using data mining. *2014 First International Conference on Networks & Soft Computing (ICNSC2014)*.
DOI : 10.1109/CNSC.2014.6906719
- [15] J. H. Park, G. H. Cha, K. H. Kim, D. C. Lee, K. J. Son & J. Y. Kim. (2015). Implementation of Crime Prediction Algorithm based on Crime Influential Factors. *Journal of Satellite, Information and Communications*, 10(2), 40-45.

김 원(Won Kim)

[정회원]



- 1990년 2월 : 한국항공대학교 항공전자공학과(공학사)
- 1999년 2월 : 한국과학기술원 전기및전자공학과(공학석사)
- 2007년 8월 : 한국과학기술원 전자전산학과(공학박사)
- 1990년 3월 ~ 1997년 2월 : 국방과학연구소 연구원
- 2000년 3월 ~ 2007년 7월 : 우송공업대학 디지털전자정보계열 교수
- 2007년 8월 ~ 현재 : 우송대학교 IT융합학부 교수
- 관심분야 : 비전 시스템, 무선통신, 로봇틱스, 유전자 알고리즘
- E-Mail : kimwon@wsu.ac.kr