

지능형 스피커 중심의 서비스 로봇 시장 분석

이성훈¹, 이동우^{2*}

¹백석대학교 ICT학부 교수, ²우송대학교 컴퓨터정보학과 교수

Analysis on Service Robot Market based on Intelligent Speaker

Seong-Hoon Lee¹, Dong-Woo Lee^{2*}

¹Professor, Division of ICT, Baekseok University

²Professor, Department of Computer Information, Woosong University

요 약 현재 우리 사회에서 자주 언급되고 있는 단어들 중 하나는 스마트 머신이라 할 수 있다. 스마트 머신은 똑똑하거나 혹은 지능적인 기능을 포함한 기계를 의미한다. 이러한 스마트 머신이 최근 우리 가정환경에서 적용되고 있는 사례들이 늘어나고 있다. 바로 스마트 홈의 결과로서 발생하는 현상들이다. 스마트 홈 환경에서 스마트 스피커는 전통적인 음악 재생의 기능에서 벗어나 이제는 스마트 홈의 다양한 구성 요소인 디바이스들을 제어하는 인터페이스 역할이 증대되고 있다. 본 연구에서는 국내, 외 스마트 스피커 시장의 기술 동향 등을 살펴보고, 현 제품들에 대한 문제점을 분석하였으며, 필요한 핵심기술에 대해 기술하였다. 국내 스마트 스피커 시장에서는 SKT와 KT가 관련 산업을 리드하고 있으며, 반면에 국외에서는 대표적인 IT 기업인 아마존, 구글, 애플 등이 관련 제품의 출시 및 기술 개발에 주력하고 있다.

주제어 : 스마트 머신, 스피커, 음성인식, 사물인터넷, 지능, 서비스 로봇.

Abstract One of the words frequently mentioned in our society today is the smart machine. Smart machines are machines that contain smart or intelligent functions. These smart machines have recently been applied in our home environment. These are phenomena that occur as a result of smart home. In a smart home environment, smart speakers have moved away from traditional music playback functions and are now increasingly serving as interfaces to control devices, the various components of a smart home. In this study, the technology trends of domestic and foreign smart speaker market are examined, problems of current products are analyzed, and necessary core technologies are described. In the domestic smart speaker market, SKT and KT are leading the related industries, while major IT companies such as Amazon, Google and Apple are focusing on launching related products and technology development.

Key Words : Smart machine, Speaker, Voice recognition, IoT, Intelligence, Service robot.

1. 서론

스마트 머신은 최근 우리 사회의 다양한 매체에 나타나는 핵심 키워드로서 우리 사회의 모든 영역에서 사용되고 있다. 스마트 머신은 똑똑한 기계라고 할 수 있으며, 여기에는 '똑똑함'과 '기계'라는 두 단어의 의미가 포함되어 있다. 이러한

스마트 머신은 다양한 분야에서 연구되고 있으며, 일부는 현재 개발되어 우리 생활속에서 이용되고 있는 사례들이 증가하고 있다. 대표적인 스마트 머신으로는 자동차 산업을 대신하고 있는 스마트 자동차인 자율자동차를 비롯하여 드론, 청소기, 스피커등이 있다. 또한 의료산업과 관련해서는 IBM의 왓슨을 들 수 있다. 애플의 시리 (Siri)는 소리를 이용하여 질문을 하고 질문에 대한 결과를 음성으로 나타내는 인식 기반

*Corresponding Author : Dong-Woo Lee(dwlee@wsu.ac.kr)

Received March 29, 2019

Accepted May 20, 2019

Revised May 3, 2019

Published May 28, 2019

의 대표적인 스마트 머신이라 할 수 있다.

로봇은 사람과 나란히 작동하는 기기이다. 기본 작업 동작을 수동으로 가르치면 주변의 상황을 파악하면서 사람들과 충돌하지 않도록 자율적으로 작업을 수행한다.

본 논문에서는 스마트 기기 중 스마트 홈 시대의 대표적인 서비스 로봇 중 하나인 스마트 스피커의 대표적인 사례를 중심으로 현재의 연구 동향과 문제점 및 발전 전망 등을 기술하였다.

사용자 인터페이스(user interface)는 사용자가 머신과의 상호 작용(communication)을 용이하게 할 수 있도록 도와주는 조정자 역할을 수행한다. 사용자 인터페이스는 개인용 컴퓨터인 PC 시대를 맞이하여 개발된 마우스 및 그래픽을 중심으로 한 사용자 인터페이스(User Interface) 플랫폼이 현재까지 이용되고 있는 중이다.

새로운 기술을 바탕으로 새로운 기기들이 만들어지고 이용되면서 사용자 인터페이스 환경도 변화를 가져오기 시작하였다. 인터페이스 환경 변화의 핵심으로는 스마트 폰을 들 수 있다. 대부분의 스마트 폰에서 사용되고 있는 사용자 인터페이스 방식은 마우스나 그래픽 환경이 아닌 사용자가 화면을 클릭하여 서비스를 이용할 수 있는 터치 기반의 방식으로 변화되었다. 스마트 폰에서의 주요 인터페이스 수단으로 터치 방식을 채택하여 사용되고 있지만 이와 더불어 최근에 관심을 갖고 개발되고 있는 차세대 인터페이스 방식으로는 음성에 기반 한 방식에 관심을 기울이고 있다.

음성 기반의 인터페이스 방식은 IoT (Internet of Things) 시장의 한 부분인 스마트 홈(Smart home) 분야에 가장 직접적으로 적용될 수 있다. 따라서 관련 기업들의 스마트 스피커에 대한 관심이 높아지고 있는 것이다. 스마트 홈은 사물 인터넷을 이용하여 다양한 가전기기를 비롯하여 수도, 전기 사용량 등을 모니터링하고 제어 할 수 있는 융/복합 환경을 의미한다. 융/복합 산업의 한 분야인 스마트 홈 환경에서 사용되는 다양한 기기들을 제어하기 위한 사용자 인터페이스로서 손을 이용하는 현재의 터치 방식보다는 음성을 이용하여 원하는 기기의 제어 및 관리가 가능하게 되었다. 이것이 바로 우리 생활환경에서 인공지능 기반의 스마트 스피커가 필요한 이유 중 하나일 것이다.

Table 1에서 기술된 2017~ 2018년도 기술 변화에 대한 가트너의 10대 전략 기술에 관한 발표 자료를 살펴보면 핵심적인 의미를 지닌 단어로 'Intelligent'를 들 수 있다.

본 연구에서는 'intelligent'의 의미를 포함하고 있는 다양한 디바이스 중에서 지능형 스피커에 대해 기술하였다. 지능형 스피커가 사용자 인터페이스로 대체되고 있는 현재의 상황 배경 및 이러한 흐름에 따라 관련 기업에서 이루어지고 있는 연구 및 관련 제품 개발 현황 및 현재 사용되고 있는 제품에서 나타나고 있는 문제점들을 알아보았다. 또한 지능형 스피커가 포함해야 하는 필수적인 기술요소들을 기술하였다. 마지막으로 지능형 스피커를 포함하는 서비스 로봇 시장의 발전 전망을 기술하였다.

Table 1. 2017-2018 Technology Trend

Category	2017 Strategy technology	2018 Strategy technology	Remarks
Intelligent	Applied AI & Advanced Machine Learning	AI Foundation	Artificial intelligence enhancement
	Intelligent Apps	Intelligent Apps	Intelligent application
	Intelligent Things	Intelligent Things	Intelligent devices
Digital	VR & AR	Digital Twins	
	Digital Twins	Cloud to the Edge	Edge computing
	Block chain & Distributed Ledgers	Conversational Platform	Interactive platform
	Conversational System	Immersive Experience	- VR, MR
Mesh	Mesh App & Service architecture	Block Chain	
	Digital Technology Paltforms	Event-Diven Model	Case-oriented model
	Adaptive Security Architecture	Continuous Adaptive Risk and Trust	continually increasing security

2. 지능형 스피커 현재 상황

2.1 지능형 스피커의 배경

지능형 스피커 등장的主된 배경은 융/복합 산업의 등장으로 산업의 한 축으로 성장한 스마트 홈 분야에서, 각 가정에서 다루어지는 다양한 기기들을 효율적이고 편리하게 제어, 관리할 필요성에서 시작되었다고 할 수 있다. 초기에 마우스 및 그래픽 환경으로 기기와의 상호작용이 이루어졌으며, 스마트 폰이 등장하면서 손을 이용한 터치스크린 패턴으로 인터페이스가 이루어지고 있다. 현재에 이르기까지의 사용자 인터페이스 분야의 변이 과정을 정리하면 다음 Table 2와 같다.

Table 2. User Interface Transition Process

	Interface method
PC age	- Mouse, Graphics
Smartphone	- Touch screen with hand
Smart Home	- Using Speech Recognition Speakers

2.2 시장 현황

미국 시장에서의 지능형 스피커 분야를 리드하고 있는 기업군들을 보면 ICT와 연관된 거대 기업들이 현재 시장을 선점하고 있다고 할 수 있다.

아마존 닷컴(Amazon)에서는 에코(Amazon Echo, 간단히 에코)라 명명된 스피커를 개발하였다. 이 장치의 외형적인 모습은 원기둥 형태를 보이고 있으며 7피스 빔포밍을 갖추고 있다[1]. 아마존 에코는 알렉사에 연결되어 사용되고 있다[2,3]. 에코를 이용하는 사용자들은 다양한 형태의 부가 서비스들을 이용할 수 있으며 주요 활용 기능으로는 음악 관련 서비스를 비롯하여 알람 설정, 날씨, 트래픽 및 실시간 정보 제공 등이 있다.

구글(Google)에서는 구글 홈으로 명명한 음성 인식 지능형 스피커를 개발하였다. 구글 홈(Google Home)은 구글 어시스턴트를 탑재하여 사용자의 음성 명령에 따라 다양한 부가 서비스를 제공한다. 서비스 형태로는 - 지식검색, 음악 재생, 가전제품 제어 등 -의 기능을 제공한다.

애플에서 개발한 지능형 스피커로는 홈팟(HomePod)이 있다. 스피커의 주요 서비스 기능은 아마존 에코와 구글 홈과 비슷하다[4]. 홈팟은 애플 A8 칩을 사용하고 시리 소프트웨어를 이용한다[5].

국내의 지능형 스피커 시장을 주도하는 기업으로는 SK

텔레콤과 KT가 있다[6,7].

먼저 SK 텔레콤에서는 누구(NUGU)로 명명된 지능형 스피커를 개발하였다. 이 스피커는 2016년 9월에 출시되었으며, 국내 최초의 지능형 스피커라 할 수 있다[8]. 2017년 8월에 NUGU mini 스피커(NU200)를 출시하였고, 또한 2018년에는 누구 캔들(NUGU candle, NU110)이 출시되었다. 2017년 8월에는 이동형 인공지능 기기인 ‘누구(NUGU) 미니’를 출시하였다. ‘누구 미니’는 기존의 스피커인 ‘누구’와 같은 기능들을 제공하면서도, 가격측면에서는 낮은 가격, 크기면에서도 더욱 소형화되었다. 유선 기반의 AI 기기가 갖고 있던 장소 제한성을 극복했다. 기존의 누구 스피커가 지원하는 대부분의 기능들을 ‘누구 미니’에서도 지원하고 있다.

KT에서는 ‘기가지니’라 명명된 지능형 스피커를 출시하였다. 2017년 1월 출시된 기가지니는 세계 최초로 IPTV와 인공지능의 융합으로 TV와 연계한 홈 비서 기능을 제공한다. 이 스피커의 주된 기능은 지능형 스피커가 아니라 엄연히 셋탑 박스다[6].

기가지니가 다른 제품과 가장 다른 점은 TV와 연결되어 화면을 이용하여 대화가 가능하다는 점이다. 하지만 지능적인 면에서는 다른 회사의 스피커보다는 학습 효율이 떨어진다는 점이다. 지능형 스피커의 주요 특성은 아래의 Table 3과 같다.

Table 3. Major Speaker Special Contents

Speaket name	Major Specialties
Alexa	- Specialized in intelligent orders
Google Home	- Specialized in Google service connection
HomePot	- Siri service connection
GigaGini	- Specialized in set-top box connection

2.3 지능형 스피커 활용 분석 및 문제점

음성인식 스피커를 사용하는 이용자 300명을 대상으로 한국 소비자 보호원에서 온라인 설문조사를 실시하였다 [9-11]. 음성인식 스피커 이용실태 조사 결과에서 추출된 주요 사항은 다음과 같다. 먼저, 제품 구매동기 항목에 대한 결과는 아래 Table 4와 같다.

Table 4. Purchase Motivation

Purchase motivation item	Percentage(%)
AI(Artificial intelligence)	67.7
Acquaintance introduction	20
Speech recognition function	6
Others	6.3

다음으로 제품 사용기간에 대한 조사 결과 내용은 아래 Table 5와 같다.

Table 5. Product Use Period

Product use period	Percentage(%)
Less than 1 month	36.3
Less than 1 ~ 3 months	43.7
Less than 3 ~ 6 months	15.4
Others	4.6

음성인식 스피커를 사용하면서 주로 활용하는 기능을 조사한 결과는 다음 Table 6과 같다. 각 내용에 대한 결과는 중복 응답을 허용한 결과이다.

Table 6. Utilization Function

Major utilization function	Percentage(%)
Play music	71.3
Weather Traffic Information	41.0
Internet Information Retrieval	40.3
Timer, schedule management	35.7

마지막으로 음성인식 스피커와 관련된 주요 불만 사항으로는 아래 Table 7과 같다. 이같은 내용을 종합하여 정리하면 다음과 같은 결론을 얻을 수가 있다.

현재 지능형 스피커를 주로 사용하는 이유는 지능형에 대한 기대감이 작용한 것임을 추측할 수 있다. 또한 지능형에 대한 기대감으로 구입을 하였으나 이용후의 결론은 기대만큼의 지능성을 보장하지 못한다는 것이다. 이같은 문제점은 크게 3가지를 들 수 있다. 먼저 소리가 불분명하거나 빠른 경우에 대한 인식을 저하의 문제가 해당된다. 둘째로 사용자의 음성이 아닌 주변의 다른 사람의 음성을 명령으로 오인식

하는 부분이 해당된다. 마지막으로 사용자와 스피커와의 자연스러운대화가 이루어지려면 먼저 스피커가 사용자에게 대화를 시작할 수 있는 환경이 이루어져야 한다. 하지만 지금의 스마트 홈 환경에서는 불가능하므로 이에 대한 노력이 이루어져야 할 것이다.

Table 7. Contents of Main Complaint

Main Content	Details of Complaint	Percentage(%)
Voice command	- If the pronunciation is incorrect and fast, the command can not be processed clearly.	56.7
Natural conversation	The speaker did not try to talk to the user firstly	45.7
External noise errors	- Recognize external noise as a user command and perform a malfunction.	37

3. 지능형 스피커의 요소 기술

스마트 스피커의 핵심은 음성인식과 사물인터넷이라 할 수 있다. 사용자들이 스피커를 통해 희망 사항 혹은 명령을 전달하면 이를 인식하여 해당 사항들에 대해 서비스를 수행하기 때문이다.

패턴인식의 한 분야인 음성인식 기술은 사용자의 음성을 입력받은 후에 컴퓨터가 해당 음성을 분석하고 특징을 추출한다[12]. 그런 다음 추출한 특징점 들을 미리 수집한 음성 모델 DB(데이터베이스)와의 유사도를 측정하며, 그 결과가 가장 근사한 것을 문자 혹은 명령어로 변환한다. 스마트 스피커 분야에서 이용되는 음성인식 과정은 크게 전 처리 과정과 인식과정으로 처리된다. 전 처리 과정에서는 이용자의 음성으로부터 인식 대상이 되는 구간을 찾는 다음 잡음성분을 제거한 후에 인식 과정을 위한 특징을 추출한다. 인식 과정에서는 입력된 음성을 음성 데이터베이스와의 비교/분석을 하여 가장 근접한 단어를 인식결과로 출력한다.

음성인식이라는 인터페이스를 통해 다양한 기기들을 제어할 수 있기 위해서는 사물인터넷 기술이 필요하다. 사물인터넷은 인터넷을 이용하여 다양한 종류의 기기들이 상호작용을 통하여 부가적인 서비스를 수행할 수 있는 기술이다. 음성인식을 위한 시스템 구조는 아래 Fig. 1과 같다.

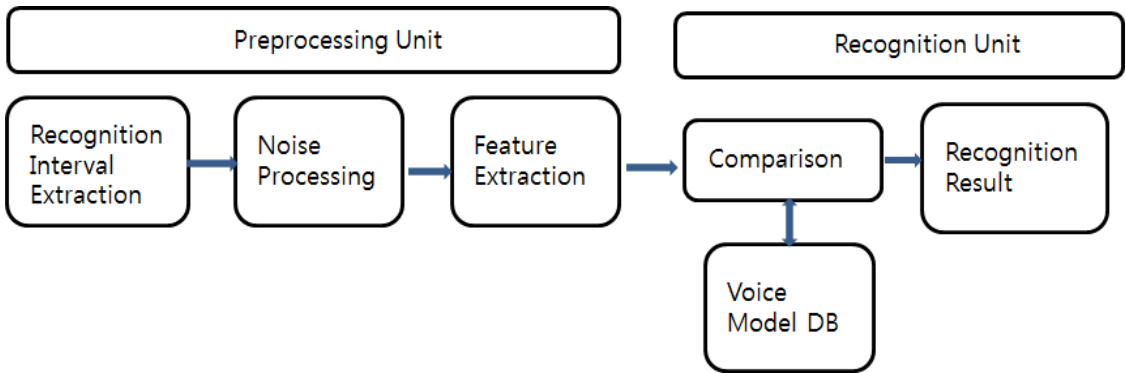


Fig. 1. Voice Recognition Procedure

사물인터넷 기술은 통신과 센서 및 프로세서와 같은 물리적인 기기 및 기술과 함께 이를 제어할 인터페이스가 합쳐진 기술이다.

4. 지능형 스피커의 향후 전망

지능형 스피커 시장과 관련하여 우리의 생활 환경이 스마트 홈 환경으로의 확대가 일반적인 추세라 할 수 있다. 이 같은 생활환경의 패턴 변화는 다양한 방면에서의 변화를 이끌 것이다. 지능형 스피커 환경도 점점 더 확대할 것이며, 사용자 만족도를 높이기 위한 관련 기술 개발에도 변화를 가져올 것이다.

지능형 스피커의 확대에 대한 향후 전망을 나타내는 미국의 사례들을 보면 향후 스피커 시장과 관련된 발전은 점점 적용 범위가 넓어질 것이며, 발전 속도 또한 빨라질 것이다. 어쩌면 스마트 홈 환경이 점점 더 확대될 가능성도 있다 [13,14]. 가트너는 Inside Radio에 따르면 미국 가구의 75%가 2020년까지 아마존 에코 (Amazon Echo) 또는 구글 홈 (Google Home)과 같은 똑똑한 스피커를 갖게 될 것이라고 예측하고 있다. 올해 초 에디슨 리서치 (Edison Research) 조사 자료에 따르면 현재 미국 가정의 7%가 보이스보더가 내장 된 스마트 스피커를 갖고 있는 것으로 보인다.

YCHARTS가 보고 한 미국 센서스 국 데이터에 따르면 2016 년 말 미국 가정의 수는 1억 1,850 만 명이다. 5년 성장률은 매년 1.48%로서, 그 추세가 계속된다면 Fig. 2에서처럼 2020년 말에 1억 2380만 미국 가정이 생길 것으로 예측했다. 가트너의 모델은 2020 년에 가정의 20%가 2개의 기기를, 5%는 3개 이상을 보유 할 것으로 예측했다.

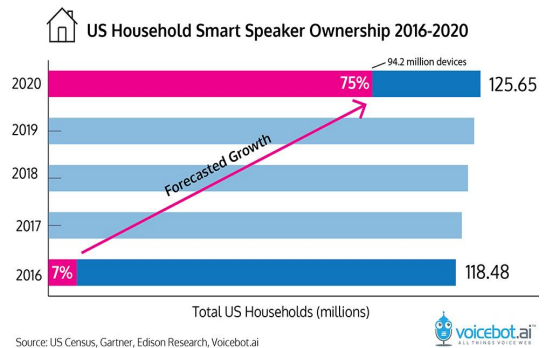


Fig. 2. Smart Speaker Ownership 2016~2020

또한 가트너와 Statista의 Digital Market Outlook에 따르면, VPA 지원 무선 스피커 시장은 2020년까지 20억 달러 규모의 비즈니스가 될 예정이며 Fig. 3에서처럼 홈 오토메이션 매출의 약 10%를 차지할 것으로 예상된다[14].

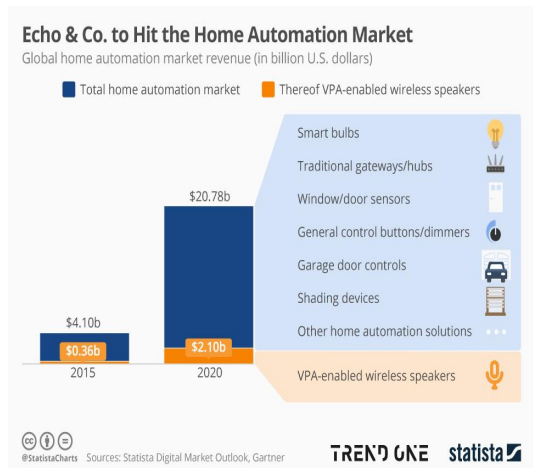


Fig. 3. Speaker Market Forecast

5. 결론

스마트 스피커는 스마트 머신의 한 유형으로 현명한 자에 속하는 동시에 생활 로봇으로도 분류된다. 본 연구에서는 스마트 스피커에 대해 국내 및 국외의 관련 기업들의 개발 사례들을 기술하였다. 국외 사례의 대표적인 개발 사례로는 아마존의 알렉사를 비롯하여, 구글의 홈, 애플의 홈팟등이 있다. 반면에 국내에서의 대표적인 사례로는 SKT NUGU와 KT의 Giga Gini가 출시되었다.

현재 이용되고 있는 스피커에 대한 실태조사를 통해 이 사용자들이 느끼는 불편 사항을 알아보았다. 대부분의 사용자들은 음성 인식 스피커가 음성을 인식하여 상호작용을 한다는 측면에서 인식율에 대해 불만을 나타내는 것으로 보인다.

스마트 스피커는 스피커의 기능뿐만이 아닌 다양한 기능을 갖고 있다. 이같은 이유는 현재의 스피커의 용도가 기존의 음악 등을 듣는 스피커 고유의 기능보다는 스마트 홈 사회에서의 다양한 홈 기기들을 음성으로 제어하는 등의 사용의 편리성 및 다양성에 주목되어 감에 따라 스마트 스피커의 중요성은 점점 더 확대되어 질 것으로 보인다.

REFERENCES

[1] B. Stone & S. Sopre. (2014), *Amazon Unveils a Listening Talking Music-Playing Speaker for Your Home*. Bloomberg, <https://www.bloomberg.com/news/articles/2014-11-06/amazon-echo-is-a-listening-talking-music-playing-speaker-for-your-home>

[2] Amazon Device Support, Set Up Your Amazon Echo, <https://www.amazon.com/gp/help/customer/display.html?nodeId=201601770>

[3] D. Bohn. (2017). *You can finally 'computer' to your echo to command it*, *The Verge*.

[4] C. Gartenberg. (2017). *Apple announces HomePod speaker to take on Sonos*, *The Verge*.

[5] M. Singleton. (2017). *Apple's HomePod delayed next year*, *The Verge*.

[6] J. Bong. (2018). *Pros and cons of KT that evolved from set-top box to artificial intelligence*, *Daily PoP*.

[7] <https://namu.wiki/w/>

[8] <https://namu.wiki/w/NUGU>

[9] M. Y. Kang. (2017). *Artificial Intelligence (AI) Home Appliances Problems and Improvements (Focusing on Speech Recognition Speakers)*, *Korea Consumer Agency*.

[10] *Artificial intelligence (AI) speakers, high customer*

satisfaction Need to improve voice recognition and conversation function.

http://www.kca.go.kr/brd/m_32/view.do?seq=2305,

[11] *Artificial intelligence speakers, 'everyday conversation' function should be further improved.* https://www.kca.go.kr/brd/m_47/down.do?brd_id=P001&seq=9902&data_tp=A&file_seq=1,

[12] J. W. Kim, S. J. Park, G. Y. Min & K. M. Lee. (2017). Virtual Reality based Situation Immersive English Dialogue Learning System. *Journal of Convergence for Information Technology*, 7(6), 245-251. DOI : 10.22156/cs4smb.2017.7.1.001

[13] B. Kinsella. (2017). *Gartner Predicts 75% of US Households will have Smart Speakers of 2020, Giving Voice to a Revolution.*

[14] Felix Richter. (2016). *Echo&Co. to Hit the Home Automation Market*, *Statistics Portal*.

이 성 훈(Seong-Hoon Lee)

[정회원]



- 1995년 2월 : 고려대학교 일반대학원 컴퓨터학과 (이학석사)
- 1998년 2월 : 고려대학교 일반대학원 컴퓨터학과 (이학박사)
- 1998년 3월 ~ 현재 : 백석대학교 정보통신학부 교수

- 관심분야 : 분산시스템, 웹서비스, 그리드 시스템, 컨버전스, 융합산업등
- E-mail : shlee@bu.ac.kr

이 동 우(Dong-Woo Lee)

[정회원]



- 1984년 8월 : 고려대학교 일반대학원 컴퓨터공학 (공학석사)
- 2005년 2월 : 고려대학교 일반대학원 전산학과 (이학박사)
- 1995년 3월 ~ 현재 : 우송대학교 컴퓨터정보학과 교수

- 관심분야 : 웹기반분산시스템, 능동시스템, 데이터베이스, 컨버전스등
- E-mail : dwlee@wsu.ac.kr