한국정보전자통신기술학회논문지(jkiiect)'16-12, Vol.9 No.6
논문 16-09-06-526 http://dx.doi.org/10.17661/jkiiect.2016.9.6.526

개인의 위치를 보호하기 위한 효율적인 더미 생성

채천원*, 송두희*, 윤지혜*, 이원규* 김용갑*, 박광진**

Efficient Dummy Generation for Protecting Location Privacy

Tian-Yuan Cai*, Doo-Hee Song*, Ji-Hye Youn*, Won-Gyu Lee*,

Yong-Kab Kim*, Kwang-Jin Park**

요 약 위치기반서비스(LBS; location based services)에서 사용자의 위치를 보호하는 연구가 많은 관심을 받아오고 있다. 특히 k-익명화(k-anonymity)를 이용한 연구가 가장 인기 있는 사생활 보호 기법이다. k-익명화란 k-1개의 다른 더미(dummy) 또는 클라이언트를 선택하여 클로킹 영역을 계산하는 기법을 말한다. 질의자는 신뢰할 수 없는 서버 또는 공격자에게 1/k의 확률로 자신의 위치 노출 확률을 줄일 수 있다. 그러나 더미가 사용자 주변에 밀집되어 있거나 질의자가 존재할 수 없는 곳에 더미가 생성된다면 질의자의 위치가 공격자에게 노출될 수 있다. 따라서 본 논문에서 우리는 실제 도로환경을 고려해서 더미를 생성함으로서 사용자의 위치보호를 높일 수 있는 시스템 모델과 알고리즘을 제안한다. 실험 결과를 통하여 제안기법의 우수성을 증명하였다.

Abstract The researches protecting user's location in location-based services(LBS) have received much attention. Especially k-anonymity is the most popular privacy preservation method. k-anonymization means that it selects k-1 other dummies or clients to make the cloaking region. This reduced the probability of the query issuer's location being exposed to untrusted parties to 1/k. But query's location may expose to adversary when k-1 dummies are concentrated in query's location or there is dummy in where query can not exist. Therefore, we proposed the dummy system model and algorithm taking the real environment into account to protect user's location privacy. And we proved the efficiency of our method in terms of experiment result.

Key Words : cloaked region, dummy, k-anonymity, location based services, protecting location privacy

1. 서론

무선통신기술과 위치측위기술의 발전으로 이동객체는 GPS(Global Positioning System)를 통해 정확한 위치 정보를 얻을 수 있다. 또한 이러한 위치측위기술의 발달로 위치기반 서비스(Location-Based Services 이하 LBS)의 이용이 확대되고 있다. LBS는 사용자의 위치정보를 LBS 서버에게 보내서 사용자가 원하는 메시지와 서비스를 제공 받을 수 있다[1-3].

LBS의 예는 지도검색(예: Google maps), 주변에 있는 PoI(Point of Interest)를 찾는 앱(예: Aroundme), 쇼핑몰의 쿠폰이나 할인 정보를 주는 앱(예: Groupon), GPS(TomTom)와 위치를 활용한 서비스(예: foursquare) 등이 있다. 공간 질의는 크게 두 가지로 분류할 수 있다. 그 중 첫 번째 스냅샷(snapshot) 질의는 사용자의 위치에서 1km 내에 있는 식당의 위치를 검색하는 질의이고, 두 번째 연속적 (continuous) 질의는 여러 개의 스냅샷 결

*Department of Information and Communication Engineering, Wonkwang University
** Corresponding Author : Department of Information and Communication Engineering, Wonkwang University (kjpark@wku.ac.kr)
Received September 09, 2016 Revised October 21, 2016 Accepted November 28, 2016

과를 연속적으로 구성한 것이다. 예를 들어 사용자가 서버에게 이동하면서 연속적으로 질의를 요청하면 서버는 질의자 위치에 따라 변화된 결과를 사용자에게 반환한다.

최초의 LBS는 군사적 목적으로 활용되었지만, 현재는 다양한 영역에서 활용되고 있다. LBS는 편리함 외에 사생활 침해의 문제를 안고 있다. 사용자는 LBS정보를 얻기 위해서 자신의 위치를 LBS 서버에게 보내게 되는데 만약 LBS 서버가 사용자의 위치정보를 악의적인 제3자에게 노출시킨다면 사용자의 취미, 이동패턴 등을 파악해서 다양한 범죄에 활용될 수 있다. 따라서 사용자의 개인 사생활 정보를 보호하는 것이 중요하다[4].

질의자의 사생활을 보호하기 위해서 다양한 연구가 제안됐다[5-10]. 시스템 환경에 따라서 두 종류로 나눌 수 있다. 1) middleware-based approach[5-7], 2) client-based approach[8-9, 14]. 첫 번째 시스템 환경은 미들웨어(middleware: e.g., trusted third party, anonymizer server)가 질의자에 대한 위치와 질의 내용을 모두 관리한다. 두 번째 시스템 환경은 미들웨어가 없다고 가정하고, 질의자는 스스로 자신의 위치를 익명화하고 LBS 서버에게 질의를 요청한다. 왜냐하면 미들웨어 또한 잠정적인 공격자가 될 수 있기 때문이다[8, 15]. 우리는 두 번째 시스템 모델에서 가짜 사용자(dummy user)를 이용한 익명화 기법을 선택한다. 더미란 질의자가 서버로부터 질의자의 위치 정보 등을 보호하기 위해 만든 가상의 질의자이다.

그러나 질의자가 잘못된 더미를 생성할 경우 사람이 존재할 수 없는 곳에 더미가 생성될 수 있다. 따라서 질의자가 생성한 k-1개의 더미를 만족하지 못하게 되어 질의자의 위치가 노출될 확률이 증가하게 된다. 우리는 이러한 문제를 해결하기 위해서 실제 환경의 장애물을 고려한 더미 생성 기법(ODG: Obstacle-based Dummy Generation)을 제안한다.

우리의 주요 기여를 정리하면 다음과 같다.

- 우리는 악의적인 미들웨어로부터 질의자의 프라이버시를 보호한다.

- 우리는 이전 기법[15-16]과는 달리 실제 존재하는 질의자의 실제환경을 고려해서 더미를 생성하기 때문에 질의자의 위치가 노출될 확률을 줄일 수 있다.

- 우리는 실험을 통해 질의자의 위치 정보 보호의 우수성을 증명했다.

2장에서는 사용자의 프라이버시를 보호하는 관련연구를 소개하고, 3장에서는 우리가 제안한 ODG에 대해 설명한다. 4장에서는 실험결과를 보여주고, 5장에서 결론을 내린다.

2. 관련연구

질의자는 프라이버시를 보호하면서 질의자가 원하는 질의 결과를 얻기를 희망한다. 질의자의 위치를 보호하는 방법은 (a) 공간 숨김 (spatial cloaking), (b) 모호화 (obfuscation), (c) 더미를 이용한 익명화(dummy-based anonymization)로 분류된다[12-13, 17-21].

공간 숨김(spatial cloaking)은 질의자의 정확한 위치를 미들웨어에게 전송하면 미들웨어는 질의자의 질의 내용을 확인한 후 서버에게 클로킹 영역과 질의 내용을 전송한다[8-10]. 클로킹 영역에 포함된 k(질의자 포함)개의 클라이언트를 구성함으로 1/k의 확률로 질의자의 위치와 질의 내용을 보호할 수 있다. 그러나 완전히 믿을 수 있는 미들웨어가 없고, 클라이언트가 밀집한 곳에서 질의 요청 시 질의자의 위치 노출이 가능성이 증가하고, 클라이언트가 드문 곳에서 k를 만족하는 클로킹 영역을 설정할 경우 클로킹 영역이 증가하기 때문에 검색해야 할 객체가 증가하는 문제점이 발생할 수 있다.

모호화(obfuscation)는 질의자의 위치 대신 질의자 주변에 있는 교차점이나 건물의 정보를 LBS 서버에게 전송한다[10-11]. 이러한 방법을 사용하면 자신의 위치를 노출하지 않아도 되기 때문에 위치 보호에 효율적일 수 있다. 그러나 질의자 주변에 선택할 목표가 없는 경우 원거리에 있는 목표를 찾아야 하기 때문에 오차가 발생할 수 있다.

더미를 이용한 익명화(dummy-based anonymization)은 그림 1처럼 질의자는 주변에 더미를 생성한 후 자신의 위치와 더미 위치를 서버에게 보낸다. 질의자는 응답을 받아서 불필요 정보를 처리하여 가장 가까운 식당을 찾는다. 그러나 더미의 위치를 랜덤하게 생성(Random-based Dummy Generation)하면 특정한 지역에서 밀집되어 나타날 수 있다[15]. 생성된 더미의 위치가 밀집된 형태를 갖게 되면 사용자의 위치가 노출되어 질 가능성도 증가될 수 있다. 이 문제를 해결하기 위해 각도를 고려한 더미를 생성하는 기법이 제안되었다[16]. 그러나 CDG(Circle-divided Dummy Generation)는 실제 환경을 고려하지 않아서 사람이 존재할 수 없는 곳에 더미가 생성된다면 k를 만족하지 못하는 문제점이 발생한다. 그림 1은 질의자가 CDG를 이용해서 원의 호 위에 더미를 생성한 것을 보여준다. 그러나 CDG를 이용하여 더미를 생성할 경우 질의자의 위치가 노출될 확률이 높아진다. 왜냐하면 그림 1에서 더미는 총 6개지만 한강 위에서 3개 더미가 생성되기 때문에 제 3자 입장에서 쉽게 더미를 구분할 수 있기 때문이다.

따라서 우리는 질의자의 위치의 노출 확률을 줄이기 위하여 더미를 생성할 때 각도와 장애물을 고려하여 효율적인 더미를 생성하고자 한다.

3. Obstacle-based Dummy Generation

질의자는 LBS를 제공 받기 위하여 자신의 실제 위치와 자신이 만든 더미의 위치를 같이 묶어서 LBS 서버에게 보낸 후 서버는 질의자가 요청한 질의 결과를 준비해서 질의자에게 전송한다. 서버에게 질의 결과를 받은 질의자는 자기 위치를 알고 있기 때문에 불필요한 정보를 제거하고 질의자가 원하는 최종 결과를 얻을 수 있다.

그림 1. 원을 분할한 더미 생성
Fig. 1. Circle-divided Dummy Generation

그림 1은 CDG를 이용한 더미 생성 방법을 보여준다. 질의자()를 중심으로 k-1(질의자를 포함하면 k개를 만족)개의 수만큼 각도를 균일하게 분할한 후에 더미()를 생성한다. 즉, 더미의 위치가 밀집하지 않을 수 있다.

그러나 그림 1처럼 질의자가 더미를 생성할 때 실제 환경을 고려하지 않으면 질의자가 존재할 수 없는 장애물 위에 더미( : kobs)가 생성될 수 있다. kobs는 장애물(obs: obstacle) 위에 존재하는 k개의 더미를 의미한다. 그러므로 kobs의 수가 증가할수록 질의자의 위치가 노출될 확률이 증가한다. 따라서 우리는 이 문제를 해결하기 위해서 ODG를 새로이 제안한다.

그림 2. 장애물을 고려한 더미 생성
Fig. 2. Obstacle-based Dummy Generation

그림 2에서 질의자는 한강공원에 위치하여 있는 것을 보여준다. 질의자는 자신의 위치를 보호하기 위해서 익명 공간을 만들어야 한다. 그림 2에서 k는 7개로 설정한다(k는 사용자의 위치를 보호하기 위한 임의의 수이며, 그림 1, 2에서 k는 7로 가정한다). 그림 1은 각도를 통해서 원 위에서 더미들을 생성한다. 이 각도는 k의 수를 통해서 계산할 수 있다. 예를 들어 k가 7일 때 경우는 2π/6(질의자 제외) = 60°이고, 각 호 위에 랜덤으로 더미를 생성한다. ODG는 그림 2와 같이 더미를 생성하기 전에 실제 환경을 고려한 후 장애물이 존재하는 각도(Θobs)에 있는 부분(빨간 점선)을 제거하고, 남은 부분(검은 선)을 k-1개로 균일하게 분할한 후 더미를 생성한다. 따라서 기존 기법인 CDG에 비해 질의자의 위치 보호 확률을 향상시킬 수 있다.

ODG의 처리과정은 알고리즘 1과 같다.

알고리즘 1. Obstacle-based Dummy Generation

Input: 질의자가 요청한 클로킹 영역(r:반지름), k, θ_{obs}
Output: CRs[] // 도로 내에 존재하는 더미들
01: 질의자는 자신의 위치와 장애물을 확인;
02: 질의자가 원하는 k와 r을 선정;
03: **if**(클로킹 영역 (CR) 내에 장애물이 존재 시)
04: CR = CR - θ_{obs};
05: **else**
06: continue;
07: **end if**
08: result = CR / (k-1); // result는 k 등분된 CR들
09: CRs[] = result;
10: **while** (k<0)
11: CRs[] = 1;
12: **if** (CRs[x] == 1)
13: i = i - x; // CRs[x]번째 배열에 1이면 x제외
14: **end if**
15: k--;
16: **end while**
17: **return** CRs[];

4. 성가능평

4.1 실험환경

본 절에서는 ODG와 CDG의 성능을 비교했다. 실험 환경은 Intel CPU G550 2.6Ghz, memory 2GB이고, visual C++ 6.0을 이용하여 실험을 실행했다. 사용된 데이터는 100*100칸의 2차원 배열을 가정한다. 그리고 k개의 더미는 (100*100)/(k-1)칸마다 더미를 랜덤하게 배치하고[22], 장애물의 배치는 배열의 비율(%)에 따라 한 행 또는 열로 랜덤하게 배치한다. 실험 데이터 세트 값은 표 1과 같다. 장애물의 기본 설정 값은 20%이고, 더미의 수 k는 20개이다.

표 1. 실험 데이터 세트 값
Table 1. Experimental dataset values

파라미터	데이터 세트 값
배열 (칸)	100 * 100
장애물 (%)	10%, **<20%>**, 30%, 50%
k (개)	**<20>**, 30, 50, 100

4.2 실험결과

그림 3은 장애물의 비율이 증가함에 따라 사용자의 위치가 노출될 확률을 보여준다. ODG는 생성된 k를 질의자의 위치 보호에 모두 활용할 수 있지만 CDG는 장애물의 비율이 증가할수록 질의자의 위치 노출 확률이 증가하는 것으로 확인된다. 제안한 ODG는 장애지역이 클로킹 영역범위에서 제외되기 때문에 CDG에 비해 클로킹 영역이 다소 감소할 수 있다. 그러나 CDG는 질의자가 존재할 수 없는 영역을 클로킹 영역으로 포함시킬 수 있다. 따라서 사용자의 위치를 보호해주지 못하는 kobs가 생성됨에 따라 질의자의 위치가 더 쉽게 노출될 수 있다.

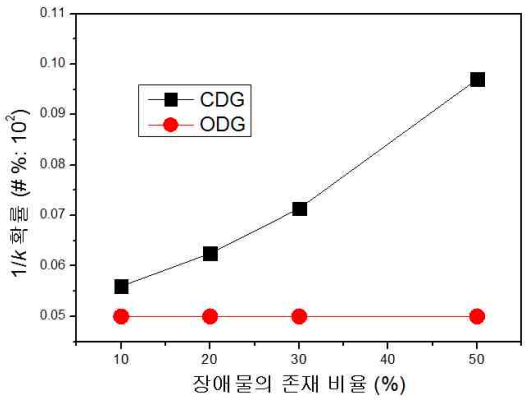

그림 3. 장애물의 존재 비율에 따른 1/k 확률
Fig. 3. 1/k probability with respect to obstacle's ratio

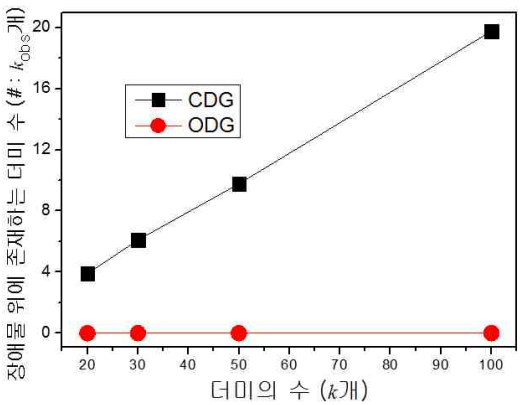

그림 4. 더미의 수(k)에 따른 장애물 위에서 존재하는 더미 수
Fig. 4. the number of dummy on obstacle(kobs) with respect to k

그림 4는 k개의 더미가 증가함에 따라 kobs가 증가하는 것을 보여준다. ODG는 장애물을 고려하기 때문에 장애물 위에 더미가 생성되지 않지만 CDG는 장애물 위에 더미가 생성되기 때문에 k개 증가할수록 kobs가 증가하는 것을 확인할 수 있다. kobs, 즉 사용자가 존재할 수 없는 위치에 생성된 더미는 사용자의 위치 노출을 증가시킬 수 있다.

5. 결론

본 논문에서 우리는 미들웨어를 제외한 시스템 모델을 제안함으로서 악의적인 미들웨어로부터 질의자의 프라이버시를 보호하고, 통신비용도 줄일 수 있었다. 또한 질의자의 주변 환경을 고려해서 더미를 생성하기 때문에 질의자의 위치가 노출될 확률을 줄일 수 있었다. 따라서 기존 CDG에 제기되었던 다양한 문제점들을 해결하였다. 끝으로 실험부분에서 제안기법의 장애물의 비율이 10%부터 50%까지 증가하면 질의자의 위치가 노출될 확률인 1/k의 값이 평균 20%인 반면에 기존기법은 28.7%가 증가했다. 이런 결과를 통해서 제안기법인 ODG의 우수성을 검증하였다.

REFERENCES

[1] K. Park and P. Valduriez, "A Hierarchical Grid Index (HGI), spatial queries in wireless data broadcasting", *Distributed and Parallel Databases*, vol. 31, no. 3, pp. 413-446, 2013.

[2] K. Park, "An Efficient Scalable Spatial Data Search for Location-Aware Mobile Services", *Information Science and Engineering*, vol. 31, no. 1, pp. 165-178, 2015.

[3] D. Song and K. Park, "A partial index for distributed broadcasting in wireless mobile networks", *Information sciences*, vol. 348, no. 20, pp. 142-152, 2016.

[4] A. Lee, S. Son, H. Kim, B. Kim, "Improving Personal Data Protection in IoT Environments", *The Journal of Korean Institute of information Security & Cyptology*, vol. 26, no. 4, pp. 995-1012, 2016.

[5] M. Gruteser and D. Grunwald, "Anonymous usage of location-based services through spatial and temporal cloaking", in *Proceedings of the international conference on Mobile systems, applications and services,* pp. 31-42, May, 2003.

[6] X. Liu, K. Liu, L. Guo, X. Li, and Y. Fang, "A game-theoretic approach for

achieving k-anonymity in location based services", in *Proceedings of the IEEE international conference on INFOCOM*. pp. 14-19, April, 2013.

[7] D. Song, J. Sim, M. Song, and K. Park, "A Privacy-Preserving Continuous Location Monitoring System for Location-Based Services", *Distributed Sensor Networks*, vol. 2015, pp. 1-10, 2015.

[8] J. Kim, E. Jeong, and B. Lee, "A design of cloaking region using dummy for privacy information protection on location-based services", *The Journal of Korean Institute of Communications and Information Sciences*, vol. 38, no. 8, pp. 929-938, 2011.

[9] C. -Y. Chow, M. F. Mokbel, and X. Liu, "A peer-to-peer spatial cloaking algorithm for anonymous location-based service", in *Proceedings of the ACM international symposium on Advances in geographic information systems*, pp. 171-178, March, 2006.

[10] M. Duckham and L. Kulik, "A formal model of obfuscation and negotiation for location privacy", in *Proceedings of the international conference on Pervasive Computing*, pp. 1-19, June, 2005.

[11] L. Šikšnys, J. R. Thomsen, S. Šaltenis, M. L. Yiu, and O. Andersen, "A location privacy aware friend locator", in *Proceedings of the international symposium on Spatial and Temporal Databases*, pp. 405-410, July, 2009.

[12] F. Olumofin, P. K. Tysowski, I. Goldberg, and U. Hengartner, "Achieving efficient query privacy for location based services", in in *Proceedings of the International Symposium on Privacy Enhancing Technologies Symposium*, pp. 93-110, July, 2010.

[13] N. Talukder and S. I. Ahamed, "Preventing multi-query attack in location-based services", in *proceedings of the ACM international conference on Wireless network security*, pp. 25-36, March, 2010.

[14] B. Niu, S. Gao, F. Li, H. Li, and Z. Lu, "Protection of location privacy in continuous LBSs against adversaries with background information", in *Proceedings of the International Conference on Computing, Networking and Communications*, pp. 1-6, February, 2016.

[15] H. Kido, Y. Yanagisawa and T. Satoh, "An anonymous communication technique using dummies for location-based services", in *Proceedings of the international conference on Pervasive Services*, pp. 88-97, July, 2005.

[16] H. Zhao, J, Wan and Z. Chen, "A Novel Dummy-Based KNN Query Anonymization Method in Mobile Services", *International Journal of Smart Home*, vol. 10, no. 6, pp. 137-154, 2016.

[17] B. Niu, Z. Zhang, X. Li, and H. Li, "Privacy-area aware dummy generation algorithms for location-based services", in *Proceedings of the IEEE international conference on Communications*, pp. 957-962, June, 2014.

[18] Y. H. Gustav, X. Wu, Y. Ren, Y. Wang, and F. Zhang, "Dummy Based Privacy Preservation in Continuous Querying Road Network Services", in *Proceedings of the IEEE international conference on Cyber-Enabled Distributed Computing and Knowledge Discovery*, pp.

94-101, October. 2014.

[19] H. Lu, C. S. Jensen and M. L. Yiu, "Pad: privacy-area aware, dummy-based location privacy in mobile services", *in Proceedings of the ACM international workshop on Data Engineering for Wireless and Mobile Access*, pp. 16-23, June, 2008.

[20] R. Kato, M. Iwata, T. Hara, A. Suzuki, X. Xie, Y. Arase, and S. Nishio, "A dummy-based anonymization method based on user trajectory with pauses", in *Proceedings of the international conference on Advances in Geographic Information Systems*, pp. 249-258, November, 2012.

[21] C. Y. Chow and M. F. Mokbel, "Enabling private continuous queries for revealed user locations", in *Proceedings of the international symposium on Spatial and Temporal Databases*, pp. 258-275, July, 2007.

[22] T. Camp, J. Boleng, and V. Davies, "A survey of mobility models for ad hoc network research", *Wireless Communications and Mobile Computing*, vol. 2, no. 5, pp. 483 - 502, 2002.

저자약력

채 천 원(Tian-Yuan Cai) [학생회원]

- 2015년 6월 : 원광대학교 정보통신학과 졸업(학사)
- 2015년 9월 ~ 현재 : 원광대학교 정보통신학과 석사과정

<관심분야> 정보통신, 위치기반서비스

송 두 희(Doo-Hee Song) [정회원]

- 2010년 2월 : 원광대학교 정보통신공학과 졸업(학사)
- 2012년 2월 : 원광대학교 정보통신공학과 졸업(석사)
- 2016년 2월 : 원광대학교 정보통신공학과 졸업(박사)
- 2016년 7월 ~ 현재 : 원광대학교 공업기술개발연구소 연구교수

<관심분야> 공간질의처리, 위치정보보호

윤 지 혜(Ji-Hye Youn) [학생회원]

- 2016년 2월 : 원광대학교 정보통신공학과 졸업(학사)
- 2016년 3월 ~ 현재 : 원광대학교 정보통신공학과 석사과정

<관심분야> 위치기반서비스, 프라이버시 보호

이 원 규(Won-Gyu Lee) [학생회원]

- 2011년 3월 ~ 현재 : 원광대학교 정보통신공학과 학사과정

<관심분야> 정보통신, 무선통신

김 용 갑(Yong-Kab Kim) [종신회원]

- 1988년 아주대학교 전자공학과 졸업(공학사)
- 1993년 앨라바마 주립대학교 (공학석사)
- 2000년 노스케롤라이나 주립대학교 (공학박사)
- 2003년~현재 원광대학교 정보통신공학과 교수
- 2006년~2013년 공과대학 POST-BK21 사업단장
- 2012년~2015년 원광대학교 창업보육센터장
- 2014년~2015년 원광대학교 창업지원단장
- 2012년~현재 LED 인력양성사업단장(전북)

<관심분야> 가시광통신시스템,광메모리센서, 전력선통신

박 광 진(Kwang-Jin Park) [정회원]

- 2000년 2월 : 고려대학교 컴퓨터학과 졸업(학사)
- 2002년 2월 : 고려대학교 컴퓨터학과 졸업(석사)
- 2006년 2월 : 고려대학교 컴퓨터학과 졸업(박사)
- 2006년 2월 ~ 2007년 2월 : 프랑스 국립컴퓨터과학연구소 (INRIA) 박사 후 연구원
- 2008년 ~ 현재 원광대학교 정보통신공학과 교수

<관심분야> 데이터베이스, 분산컴퓨팅