

서지통계학적 분석을 이용한 동형 암호의 연구경향 분석*

야마다 아키히코,^{1*} 이 은 상^{2*}¹서울대학교 빅데이터 혁신융합대학 (교수), ²세종대학교 (교수)

Analysis of Research Trends in Homomorphic Encryption Using Bibliometric Analysis*

Akihiko Yamada,^{1*} Eunsang Lee^{2*}¹Bigdata Convergence and Open Sharing System Seoul National University (Professor),
²Sejong University (Professor)

요 약

동형 암호 기술은 최근 널리 연구되고 있는 유망한 기술로서, 데이터를 암호화한 상태에서도 연산이 가능하게 하는 기술이다. 본 논문에서는 서지통계학적 분석을 통해 6,047개의 동형 암호 논문을 대상으로 연구 동향을 체계적으로 분석한다. 구체적으로 연도별 논문 수 분석, 키워드 상관관계, 주제 군집 분석, 동형 암호 관련 키워드의 연도별 변화 분석, 그리고 동형 암호 연구 수행 기관의 국가 분석을 통해 동형 암호 기술의 연구 동향을 객관적이고 정량적으로 분석한다. 이러한 분석 결과는 동형 암호를 연구하고 활용하는데 필요한 전략적인 방향성을 제공하며, 이는 후속 연구, 산업 응용 등에 큰 도움이 될 것이다.

ABSTRACT

Homomorphic encryption is a promising technology that has been extensively researched in recent years. It allows computations to be performed on encrypted data, without the need to decrypt it. In this paper, we perform bibliometric analysis to objectively and quantitatively analyze the research trends of homomorphic encryption technology using 6,047 homomorphic encryption papers from the Scopus database. Specifically, we analyze the number of papers by year, keyword co-occurrence, topic clustering, changes in related keywords over time, and country of homomorphic encryption research institutions. Our analysis results provide strategic directions for research and application of homomorphic encryption and can be a great help for subsequent research and industrial applications.

Keywords: homomorphic encryption, bibliometric analysis, cryptography, research trends, topic cluster

1. 서 론

동형 암호(homomorphic encryption)[1]는 최근 암호 분야에서 널리 연구되고 있는 암호 기술로

서, 암호화된 데이터에 대한 연산을 가능하게 하는 기술이다. 이 기술은 클라우드 컴퓨팅, 빅 데이터 분석, 사물인터넷, 인공지능 등 다양한 분야에서 활용될 수 있다.

동형 암호 기술의 발전과 활용을 위해서는 동형 암호 기술의 연구 동향을 파악하는 것이 매우 중요하다. 이미 수천 건의 논문이 존재하는 동형 암호 분야에서 본 논문은 서지통계학적 분석(bibliometric analysis)을 활용하여 동형 암호 기술의 연구 동향을 객관적이고 정량적으로 분석한다. 서지통계학적

Received(05. 11. 2023), Modified(06. 13. 2023),
Accepted(06. 13. 2023)

* 이 논문은 2023년도 정부(교육부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업임 (No. 2022R111A1A0106828412).

† 주저자, ayamada5413@snu.ac.kr

‡ 교신저자, eslee3209@sejong.ac.kr(Corresponding author)

분석은 서적, 논문, 기타 학술 자료 등의 여러 면을 정량적으로 연구하고 평가하는 방법으로, 특정 분야의 성장률, 주제 군집(topic cluster), 추세 등을 쉽고 빠르게 파악할 수 있게 한다.

본 논문에서는 Scopus 데이터베이스에서 "homomorphic encryption"이라는 검색어로 검색한 6,047개의 논문을 대상으로 VOS viewer 소프트웨어를 사용하여 다음과 같은 분석을 진행한다.

- 1) 동형 암호 연구 분야에서의 연도별 논문 수 분석을 통해 동형 암호 연구 추세를 확인한다.
- 2) 키워드 상관관계 및 주제 군집 분석을 통해 동형 암호 연구의 주요 주제들을 파악한다.
- 3) 동형 암호 관련 키워드의 연도별 변화 분석을 통해 시간의 흐름에 따른 연구 분야 추세의 변화를 확인한다.
- 4) 동형 암호 연구 수행 기관의 국가 분석을 통해 해당 분야에서 선도적인 위치를 차지하는 국가들을 파악한다.

이러한 본 논문에서의 서지통계학적 분석은 동형 암호 동향에 대한 실질적인 정보를 제공하며 동형 암호 기술의 연구와 활용을 위한 전략적인 방향성을 제공할 것이다.

II. 서지통계학적 분석

서지통계학적 분석은 서적, 논문, 기타 학술 자료 등의 다양한 측면을 정량적으로 조사하고 평가하는 방법이다. 이 분석은 통계학과 수학 모델을 활용하여 연구와 연구자들의 영향력, 생산성 및 보급을 평가한다. 분석의 핵심 목적은 특정 분야나 주제에서의 성장, 발전 및 추세에 대한 통찰력을 제공함으로써 연구자, 기관, 정책 결정자들이 데이터 기반 증거에 근거한 결정을 내릴 수 있도록 돕는 것이다. 이러한 분석은 사회학이나 역사와 같은 인문학 분야뿐만 아니라 의학[2], 생물학[3], 재료 과학[4] 등 다양한 분야에서 활용되고 있다.

서지통계학적 분석의 주요 구성 요소는 다음과 같다. 첫 번째로, 키워드 및 주제 분석이 있다. 이는 학술 문헌과 관련된 키워드와 주제를 검토하여 특정 분야의 연구 초점과 동향에 대한 통찰력을 제공한다. 분석 대상 단어는 논문의 제목과 초록에서 주로 추출된다. 두 번째로, 연구 간 관계 및 공동 연구 분석이 있다. 두 개의 출판물이 함께 인용되는 빈도를 조사

하여 서로 다른 연구 영역 또는 주제 간의 관계에 대한 통찰력을 제공한다. 또한 저자, 기관 및 국가 간의 공동 연구 패턴을 분석함으로써 연구 네트워크와 잠재적 파트너십을 식별하는 데 도움이 된다. 또한, 국가 간 연구 네트워크 분석을 통해 전 세계적인 연구 발전 현황을 파악할 수 있다. 세 번째로, 인용 분석이 있다. 다른 학술 문헌에서 인용된 문헌의 횟수를 조사하며, 이는 연구 커뮤니티 내에서 그 학술 문헌의 영향력을 나타내는 지표가 된다. 인용률이 높은 문헌은 일반적으로 우수한 연구 성과나 중요한 가치를 지닌 것으로 간주된다. 네 번째로, 영향 요인 분석이 있다. 이는 특정 분야의 저널이 받는 평균 인용 횟수를 기준으로 저널의 상대적 중요성을 평가한다. 영향력이 큰 저널은 일반적으로 권위 있고 영향력 있는 것으로 간주된다. 다섯 번째로, 출판 및 생산성 분석이 있다. 이는 특정 기간 동안 저자, 기관 또는 국가가 작성한 학술 문헌의 수와 같은 연구 결과물의 양을 평가한다. 이를 통해 주어진 분야에서 연구 활동의 활발함을 측정한다.

서지통계학적 분석은 다양한 분야의 발전에 크게 기여한다. 첫 번째로 연구 정책 결정에 도움을 준다. 정부, 기관, 대학 등이 연구 정책을 결정할 때 서지통계학적 분석을 활용하여 연구 우선순위와 연구자원 배분을 결정할 수 있다. 이를 통해 효율적인 연구 지원이 가능하며, 연구 경쟁력 향상에 기여한다. 두 번째는 기술 동향 파악에 활용된다. 특히 기업들은 서지통계학적 분석으로 기술 동향을 이해하고, 경쟁 기업들의 연구 동향을 살펴볼 수 있다. 이를 바탕으로 기술 개발 방향성을 설정하고, 시장에 적합한 제품을 개발할 수 있다. 세 번째로 학술지 평가 및 랭킹 작성에 활용된다. 서지통계학적 분석을 통해 학술지의 인용 횟수, 영향력 등을 평가하여 학술지 랭킹을 작성한다. 이를 통해 연구자들은 논문을 투고할 학술지를 결정하거나 학술지의 수준을 파악할 수 있다. 마지막으로 교육 및 연구 프로그램 개발에 기여한다. 교육 기관은 서지통계학적 분석을 활용하여 최신 경향에 맞춘 교육 및 연구 프로그램을 개선할 수 있다. 이 결과 프로그램 참여 학생들의 전문성을 향상시키고, 졸업생의 취업률을 높일 수 있다.

본 논문에서는 최근 암호 분야에서 광범위하게 연구되고 있는 동형 암호 분야에 대해 서지통계학적 분석을 실시하여 동형 암호 기술의 최근 동향을 분석하고 그 결과를 제공한다.

III. 동형 암호에 대한 서지통계학적 분석 방법

본 연구에서는 1991년부터 2023년까지 Scopus 데이터베이스에서 "homomorphic encryption"이라는 검색어로 검색한 6,047개의 논문을 대상으로 분석을 진행한다. 분석을 위해 각 논문의 제목, 초록, 키워드 등의 정보를 수집하며, 이를 바탕으로 논문 간의 연결성을 분석한다. 논문 간의 연결성은 공저자, 인용, 공동 참고문헌 등을 기반으로 계산하며, 이를 바탕으로 네트워크 분석을 수행하여 학문적인 동향을 파악하고, 학문적 네트워크 구조를 분석함으로써 해당 분야의 연구 동향과 주요 주제를 파악한다.

VOS viewer 소프트웨어[5]는 서지통계학적 분석을 수행하기 위해 널리 사용되는 도구이다. 본 연구에서는 Scopus 데이터베이스에서 수집한 6,047개의 논문 데이터를 VOS viewer를 사용하여 분석하였다. 목표는 동형 암호 분야에서 가장 중요한 연구 주제 및 협력 관계를 식별하는 것이다. 우선 Scopus에서 추출한 논문 정보 데이터를 다운로드한 후 VOS viewer로 데이터를 가져온다. 이 소프트웨어의 bibliometric mapping 및 클러스터링 알고리즘을 사용하여 키워드, 나라간의 협력관계, 인용네트워크를 시각화한다.

VOS viewer는 맵, 밀도 플롯 등 다양한 시각화 방법을 제공하여 연구자가 주요 주제 및 관계를 식별할 수 있게 한다. 본 연구의 분석은 "homomorphic encryption" 분야의 특정 하위 주제에 대한 키워드 군집뿐만 서로 다른 나라 간의 중요한 협력 관계와 인용관계도 확인한다.

VOS viewer를 이용한 서지통계학적 분석[5,6]에는 여러 용어들이 사용된다. 키워드, 국가, 출판물과 같은 개체를 나타낼 때는 "Items"라는 용어를 사용한다. "Link"는 두 item 간의 연결 또는 관계를 의미하며, 이는 item의 동시 발생(co-occurrence)를 나타낸다. "Link Strength"는 각 링크의 속성을 표현하며, 이 속성은 숫자로 나타난다. 예를 들어, 공동 저자 링크에서 Link Strength 값이 높을수록 두 연구자가 공동 저자로 활동하는 빈도가 높음을 의미한다. "Number of Links"는 item이 다른 item과 연결된 링크 수를 나타내고 "Total Link Strength"는 item이 다른 item과 가진 연결 강도의 합을 의미한다. "Documents"는 출판된 논문의 수를, "Citations"는 인용된 횟수를 나타낸다.

IV. 결 과

4.1 연도별 논문 수 분석

2010년까지 매년 100편 이하의 동형 암호 논문이 발표되었으나 2011년부터는 매년 100편 이상의 논문이 발표되어 2022년에는 무려 956편이 출판되었다. 2010년부터 논문 수가 급격하게 증가한 원인은 2009년에 Gentry에 의해 발표된 첫 완전 동형 암호 논문[7]의 영향이 크다. Gentry 논문 전까지 연구된 동형 암호는 제한된 횟수의 동형 연산만 지원했지만, Gentry 논문 발표 이후 다양한 완전 동형 암호가 개발되었고 최근에는 표준화 과정도 진행되고 있다. Fig. 1은 이러한 추세와 부합함을 보여준다.

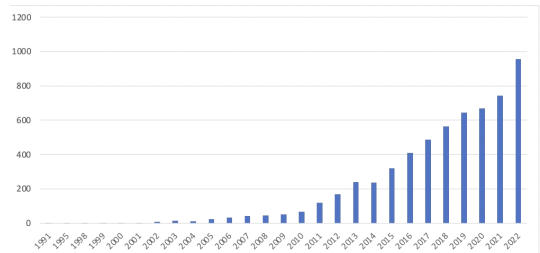


Fig. 1. Number of homomorphic encryption papers by year

4.2 키워드 상관관계 및 주제 군집 분석

Fig. 2는 "homomorphic encryption"을 키워드로 검색한 학술 문헌들에서 등장하는 키워드들의 상관관계를 분석한 것이다. 그림에서 다양한 색상은 주제 군집을 의미한다. 많이 등장하는 키워드일수록 키워드 크기와 원의 크기가 크며, 서로 상관된 키워드는 선으로 연결된다. Fig. 2에서 대표적인 5개 주제 군집을 분석해본다.

4.2.1 빨간색 주제 군집

빨간색 주제 군집은 암호 자체에 대한 이론적인 연구를 나타낸다. 동형 암호 분야와 관련된 "fully homomorphic encryption", "somewhat homomorphic encryption", 암호 자체에 대한 키워드 "cryptography", "ciphertexts", 공개키 암호에 대한 키워드 "public-key encryption" 등이 포함된다. 또한, "hardness of learning",

4.2.3 초록색 주제 군집

초록색 주제 군집은 동형 암호를 머신러닝 응용 분야에 적용하는 연구를 의미한다. “machine learning”, “deep neural networks”, “neural networks”, “machine-learning”, “convolutional neural networks”, “privacy-preserving machine learning” 등 머신러닝과 관련된 키워드들이 많이 나타난다. 더불어, 동형 암호 외에도 머신러닝의 프라이버시를 보호하는 기법 중 하나인 “differential privacy” 키워드도 확인할 수 있다. 또한, “face recognition”, “biometrics”, “biometric identification” 등 머신러닝과 관련된 구체적인 응용 키워드들이 나타난다. 결국 초록색 주제 군집은 동형 암호를 머신러닝에 적용하여 프라이버시를 보호하는 연구와 얼굴 인식, 생체 인식 등의 구체적인 응용 분야에서 동형 암호를 활용하는 연구 주제를 나타낸다.

4.2.4 노란색 주제 군집

노란색으로 표시된 주제 군집에서는 “genomics”, “genome-wise association study”, “genetic”, “gwas”, “genes”, “dna” 등 유전학에 대한 키워드가 많이 나타난다. 유전자 정보는 매우 민감한 인체 정보로, 프라이버시 보호가 매우 중요한 정보로 간주된다. 최근에는 동형 암호를 사용하여 유전자 정보의 프라이버시를 보호하려는 연구들이 진행되고 있으며, 이러한 연구들이 노란색 주제 군집에 속한다고 할 수 있다. 또한, “health care application”, “e-health”, “health care delivery”, “electronic health records” 등 건강 관리와 관련된 키워드들도 많이 나타난다. 최근 건강 관리를 원격으로 관리하는 시스템들이 많이 개발되고 있지만, 건강 정보는 민감한 개인 정보이므로 외부에 노출되는 것을 원치 않는다. 따라서, 동형 암호를 사용하여 프라이버시를 보호하면서 건강 관리를 가능하게 하는 연구[10]가 진행되고 있으며, 이러한 연구들이 노란색 주제 군집에 속한다.

4.2.5 주황색 주제 군집

주황색으로 표시된 주제 군집은 IoT와 스마트그리드 같은 분야에서 동형 암호를 적용하는 연구 주

로 볼 수 있다. 이는 이 군집에 “internet of things”, “smart power grids”, “smart meters” 등의 키워드가 포함되어 있는 것을 통해 알 수 있다. IoT와 스마트그리드에서는 다양한 기기 간 통신이 활발하게 이루어지며, 프라이버시를 보호하면서 데이터를 안전하게 전송하기 위해 동형 암호를 적용하는 연구들이 최근 진행되고 있다. 이러한 연구들이 주황색 주제 군집을 구성한다.

이처럼 동형 암호에 대한 다양한 주제 군집을 확인할 수 있으며, 이는 실제 연구 추세와 잘 부합한다. 6,000편 이상의 동형 암호 관련 논문을 모두 검토하며 추세를 분석하는 것은 쉽지 않은 일이지만, 본 논문에서는 서지통계학적 분석을 통해 신속하고도 정확하게 연구 주제 동향을 파악한다.

4.3 동형 암호 관련 키워드의 연도별 변화

Fig. 3은 시간의 흐름에 따른 키워드 네트워크 변화를 보여준다. 파란색이 짙어질수록 2016년 이전의 연관성이 높은 키워드를 나타내고, 노란색으로 갈수록 2022년의 연관성이 높은 키워드를 나타낸다. 이를 통해 연도별 사용된 키워드들의 변화 흐름을 파악할 수 있다.

파란색 키워드 중에는 “public-key encryption”, “plaintext” 등 암호 관련 기본 기술에 대한 단어들이다. 또한, “secure signal processing”, “signal processing in the encryption”과 같은 신호 처리 관련 키워드들, “sensors”, “wireless sensor network”, “ad hoc networks” 등 네트워크 관련 단어들도 눈에 띈다. 이를 통해 과거에는 신호 처리와 네트워크 분야에서 보안을 위해 동형 암호를 적용하는 연구들이 활발했음을 알 수 있다.

반면, 노란색 키워드 중에는 “machine learning”, “neural networks”, “federated learning”, “convolutional neural networks”와 같은 인공지능 관련 단어가 많이 등장한다. 이는, 최근 인공지능 서비스의 확산에 따라 인공지능에 동형 암호를 적용하는 연구들이 늘어나고 있기 때문이다.

이처럼 Fig. 3을 통해 과거와 현재의 연구 추세를 확인할 수 있으며, 노란색으로 표시된 키워드들을 중심으로 검색하면 최신 연구 추세에 맞는 후속 연구를 진행하는데 도움이 될 수 있다.

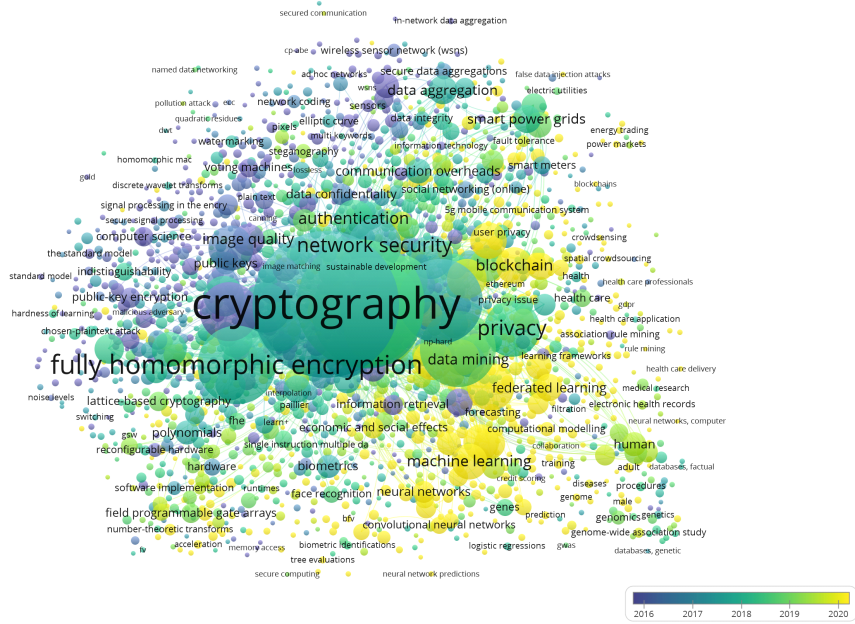


Fig. 3. Yearly changes in keywords related to homomorphic encryption

4.4 동형 암호 연구 수행 기관의 국가 분석

는데, 이는 아시아 지역에서의 연구 활동이 활발하여

Fig. 4와 Table 1은 동형 암호 연구 수행 기관의 국가를 분석한 것이다. 예상대로 미국과 중국에서 연구가 가장 활발하게 진행되고 있다. 그 다음으로 영국, 프랑스, 캐나다, 독일, 호주 순서로 연구가 이루어지고 있다. 인도, 싱가포르, 일본, 사우디아라비아와 같은 아시아 국가들 역시 순위권에 포함되어 있

Table 1. Top 15 countries in total link strength

Country	Links	Total Link Strength	Documents	Citations
United States	44	717	1,128	41,079
China	46	659	1,891	15,750
United Kingdom	40	234	214	6,427
France	36	221	277	6,776
Canada	31	201	227	7,843
Germany	33	184	267	3,790
Australia	31	180	199	3,767
India	41	152	687	3,349
Singapore	23	141	137	2,231
Japan	25	134	317	4,335
Saudi Arabia	24	125	85	1,229
South Korea	23	121	247	3,451
Switzerland	23	99	107	2,416
Israel	12	94	115	7,263
Hong Kong	17	90	86	1,771

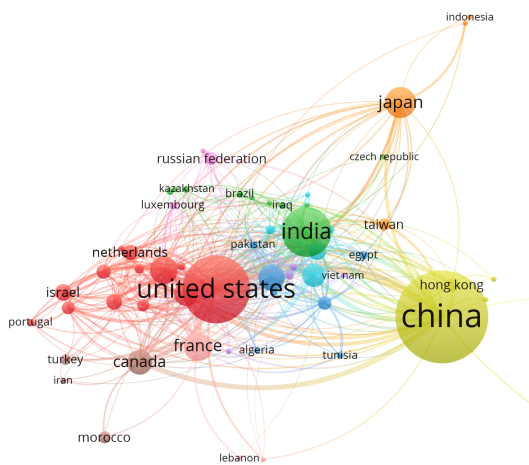


Fig. 4. National analysis of institutions performing homomorphic encryption research

동형 암호 연구에 큰 기여를 하고 있음을 보여준다.

한국은 12위에 위치해 있는데, 서울대학교 천정희 교수님 연구팀에서 최근 가장 유망한 완전 동형 암호 중 하나인 Cheon-Kim-Kim-Song(CKKS) 암호 시스템[11]을 개발한 것이 큰 기여를 한 것으로 보인다. CKKS 암호는 최근 동형 암호 표준화에서도 유력 후보로 꼽히고 있다. 한국이 CKKS 암호시스템에 대한 주도권을 계속 유지하고 연구를 확장해 나간다면, 더 높은 순위로 올라가면서 동형 암호 분야에서 전 세계적으로 큰 영향력을 발휘할 수 있을 것이다.

V. 결 론

이 논문은 서지통계학적 분석을 통해 동형 암호 분야의 동향을 체계적으로 분석한 결과를 제시하였다. 우선, 논문 수 분석을 통해 2010년 이후 동형 암호 연구가 급격히 증가한 것을 확인할 수 있었으며, 이는 Gentry 논문 발표 이후 완전 동형 암호의 다양한 발전이 이루어졌기 때문으로 보인다. 또한, 키워드 상관관계 분석을 통해 동형 암호 연구 주제 군집을 파악하였다. 동형 암호의 개발 및 안전성 분석 연구, 병렬 컴퓨팅 및 하드웨어 가속기 등을 활용한 연산 가속화, 머신러닝에의 적용, 유전 정보 프라이버시 보호, IoT 및 스마트그리드에의 적용 등의 연구 주제가 확인되었다. 그리고, 국가 분석을 통해 현재 동형 암호 연구에서 미국과 중국이 선도하고 있는 것으로 확인되었지만, 한국 역시 CKKS 암호시스템 개발 등을 통해 세계적으로 영향력을 발휘할 가능성이 있다. 본 논문에서 제시한 동형 암호에 대한 서지통계학적 분석은 후속 연구, 산업 응용, 정책 입안 등에 큰 도움이 될 것으로 기대된다.

References

- [1] J. Yoo and J. Yoon, "Analysis and trends of LWE and fully homomorphic encryption," *Review of KIISC*, 30(5), pp. 111-119, Oct. 2020.
- [2] P. Kokol, H. B. Vosner, and J. Završnik, "Application of bibliometrics in medicine: a historical bibliometrics analysis," *Health Information & Libraries Journal*, vol. 38, no. 2, pp. 125-138, Jan. 2021.
- [3] M. Abouzid, A. K. Glowka, and M. Karazniewicz-Lada, "Trend research of vitamin D receptor: Bibliometric analysis," *Health Informatics Journal*, vol. 27, no. 4, pp. 1-14, Oct. 2021.
- [4] V. H. Pauna, E. Buonocore, M. Renzi, G. F. Russo, and P. P. Franzese, "The issue of microplastics in marine ecosystems: A bibliometric network analysis," *Marine Pollution Bulletin*, vol. 149, pp. 110612, Dec. 2019.
- [5] N. J. Eck and L. Waltman, *VOSviewer Manual(2022)*, (4/23, 2023), https://www.vosviewer.com/document/Manual_VOSviewer_1.6.18.pdf
- [6] N. J. Van Eck and L. Waltman, "Visualizing bibliometric networks," *Measuring scholarly impact: Methods and practice*, pp. 285-320, Jan. 2014.
- [7] C. Gentry, "Fully homomorphic encryption using ideal lattices," *Proceedings of the Forty-First Annual ACM Symposium on Theory of Computing*, pp. 169-178, Springer, May 2009.
- [8] W. Jung, S. Kim, J. Ahn, J. Cheon, and Y. Lee, "Over 100x faster bootstrapping in fully homomorphic encryption through memory-centric optimization with GPUs," *IACR Transactions on Cryptographic Hardware and Embedded Systems*, vol. 2021, no. 1, pp. 114-148, Aug. 2021.
- [9] S. Kim, J. Kim, M. J. Kim, W. Jung, J. Kim, M. Rhu, and J. Ahn, "BTS: An accelerator for bootstrappable fully homomorphic encryption," *Proceedings of the 49th Annual International Symposium on Computer Architecture*, pp. 711-725, June 2022.
- [10] M. Kim, X. Jiang, K. Lauter, E.

Ismayilzada, and S. Shams, "Secure human action recognition by encrypted neural network inference," *Nature Communications*, vol. 13, no. 1, pp. 4799, Aug. 2022.

[11] J. Cheon, A. Kim, M. Kim, and Y. Song, "Homomorphic encryption for arithmetic of approximate numbers," *Proceedings of ASIACRYPT 2017*, LNCS 10624, pp. 409-437, Springer, Nov. 2017.

〈저자소개〉



야마다 아키히코 (Akihiko Yamada) 정회원
 2010년 3월: 일본 홋카이도대학교 농학부 학사
 2012년 3월: 일본 홋카이도대학교 국제홍보미디어관광학원 석사
 2021년 2월: 서울대학교 언어학과 박사
 2021년 3월~2022년 2월: 서울대학교 약학대학 박사후연구원
 2022년 3월~2023년 6월: 서울대학교 빅데이터 혁신융합대학 연구원
 2023년 7월~현재: 서울대학교 빅데이터 혁신융합대학 객원조교수
 <관심분야> 자연어처리, 지식그래프, 온톨로지, 머신러닝, 의미론



이 은 상 (Eunsang Lee) 정회원
 2014년 8월: 서울대학교 전기정보공학부 학사
 2020년 8월: 서울대학교 전기정보공학부 박사
 2020년 9월~2022년 8월: 서울대학교 전기정보공학부 박사후연구원
 2022년 9월~현재: 세종대학교 소프트웨어학과 조교수
 <관심분야> 동형암호, 포스트 양자 암호, 프라이버시-보호 머신러닝