Tu Hoang

POSTDOCTORAL RESEARCHER · PRIVACY · SECURITY · MACHINE LEARNING · DECENTRALIZED SYSTEMS

Jean Henry Dunant 7, Varese, Italy

📱 +39 389 142 8587 | ✉ anhtu.hoang@uninsubria.it | 🏠 tuhoang.me | 🐙 tuhoag | 💼 tuhoag | 🆔 0000-0002-1027-3905

Skills

Privacy	k-anonimity, Differential Privacy, Zero-knowledge Proof.
Cryptography	Homomorphic Encryption, Pedersen Commitment, Multi-party Computation.
Machine Learning	Federated Learning, Knowledge Graph Embedding, Graph Convolution Network, Node2Vec, Image Classification.
Blockchain	Hyperledger Fabric, Ethereum.
Programming	Python (PyTorch, Tensorflow, MXNet, Scikit-Learn), Go, Node.js, Solidity, Ruby (Rails), Bash, Scala, C#.
Databases	Redis, MongoDB, Cassandra, Neo4j.
Tools	Linux, Git, Docker, Microsoft Azure.

Work Experience

University of Insubria
Varese, Italy

Postdoctoral Researcher (The EU H020 CONCORDIA project) — *Dec 2020 - Nov 2022*

- **Privacy-preserving Deep Learning Usages of Knowledge Graphs:**
    - introduced k-**anonymity** techniques protecting data owners in anonymized knowledge graphs (KGs) (**Python**);
    - designed adapters to anonymize KGs with **Scikit-Learn**'s clustering algorithms (e.g., k-medoids, hdbscan);
    - optimized anonymized KGs to be used for **deep learning**, i.e., **Relational Graph Convolution Network** (**PyTorch**).
- **ProMark: A Transparent And Privacy-Aware Decentralized Proximity Advertising Platform:**
    - designed a **Hyperledger Fabric**-based platform to ensure advertisers', publishers', and customers' behaviors are transparent;
    - protected customers' privacy by introducing a non-interactive **zero-knowledge-proof** scheme using **Pedersen commitment**, **ECIES**, and **multi-party computation** (**Go**);
    - implemented smart contracts (**Go**) and dApps (**Node.js**) to evaluate the platform.
- **Federated Learning Platform for KG Representation:**
    - used **differential privacy** to protect the existence of users in local KGs;
    - exploited **homomorphic encryption** to aggregate local models.
- **Supervised a PhD and two Master students:**
    - *The Reward System for ProMark:* extended ProMark to calculate customers' rewards without disclosing the rewards to others.
    - *The Blockchain-Based Anonymization Platform:* designed a **zero-knowledge-proof** protocol using **Chaum-Pedersen** to allow data owners to verify whether anonymized KGs satisfy their k-**anonymity** privacy preferences by using smart contracts (**Hyperledger Fabric**).

University of Insubria
Varese, Italy

Ph.D. Student (The EU H020 CONCORDIA project) — *Oct 2017 - Dec 2020*

- **Courses:**
    - Deep Learning: learned **deep learning** tasks using **Tensorflow/PyTorch** and participated Google Landmark classification challenge hosted on Kaggle;
    - NoSQL: studied popular data stores (**Redis**, **MongoDB**, **Cassandra**, **Neo4j**), sharding/replication.
- **The Anonymization of Knowledge Graphs:**
    - introduced the state-of-the-art k-**anonymity** techniques protecting data owners in anonymized KGs in static and sequential publishing;
    - extended **node2vec (PyTorch)** to generate users' vectors such that the information loss of anonymizing two users' data is similar to the distance between their vectors;
    - developed an efficient algorithm allowing data providers to use **Scikit-Learn's** clustering algorithms to anonymize KGs.
- **Marker centroid detection in X-ray radiographs:** implemented an adapter library used for training a **ResNet-based model** to detect the centroids of metal markers in X-ray projections using **MXNet** and **Python**.

National Institute of Informatics
Tokyo, Japan

Research Intern — *Mar 2013 - Sept 2013*

- **Copyright Protection on Data Distribution:** proposed a watermarking technique to detect the illegal distribution of anonymized data (**C#**).
- **Text Anonymization:** collaborated with a Ph.D. student to detect/anonymize users' temporal locations in tweets (**machine learning** and **NLP**).

University of Science-Vietnam National University, Ho Chi Minh City (Vietnam)
Ho Chi Minh City, Vietnam

Teaching Assistant & Lecturer — *Sept 2009 - Oct 2017*

- Researched **anonymization** techniques.
- **Supervised Bachelor students** on theses collaborated with companies.
- Gave lectures on: programming (**C/C++**), fundamental of databases, Oracle security technologies, NoSQL (**Redis, MongoDB, Cassandra, sharding/replication**), and web programming (**Docker, .NET, Ruby on Rails, Node.js, Java**).

Education

University of Insubria   *Varese, Italy*
Doctor of Philosophy in Computer Science   *Oct 2017 - Dec 2020*
- **Thesis:** Privacy-preserving Publishing of Knowledge Graphs;
- **Courses:** Image Classification, Deep Learning, NoSQL.

University of Science, Vietnam National University   *Ho Chi Minh City, Vietnam*
Master in Information Systems   *Sept 2009 - Dec 2012*
- **Thesis:** Privacy-preserving Publishing of Relational Data;
- **Courses:** Digital Image Processing and Computer Vision, Distributed Systems, Knowledge Base Systems, Information Retrieval System, Database Systems, Object Oriented Information Systems, Mathematical Methods in Computer Science.

University of Science, Vietnam National University   *Ho Chi Minh City, Vietnam*
Bachelor in Information Technology   *Sept 2005 - Sept 2009*
- **Thesis:** A Turn-based Strategy Game Generation Framework for Windows Mobile;
- **Interested Courses:** Artificial Intelligence, Database Security, Introduction to Cryptography, Information Security and Applications, Data structure and Algorithms, Graph Theory, Network Application Programming, Database Management Systems, Operating Systems.

Achievements

2019	**B2**, Cambridge First B2	*Italy*
2014	**100%,** Functional Programming Principles in Scala (École Polytechnique Fédérale de Lausanne)	*Coursera*
2014	**100%,** Introduction to Data Science (University of Washington)	*Coursera*
2014	**100%,** Machine Learning (Stanford University)	*Coursera*
2014	**100%,** Introduction to Data Science in Python (University of Michigan)	*Coursera*
2009	**Top 5**, Microsoft Imagine Cup	*Vietnam*

References

Professor Elena Ferrari   *elena.ferrari@uninsubria.it*
University of Insubria   *Varese, Italy*

Professor Barbara Carminati   *barbara.carminati@uninsubria.it*
University of Insubria   *Varese, Italy*

Languages

English	Professional proficiency
Vietnamese	Native proficiency

Publications

1. **Anh-Tu Hoang**, Barbara Carminati, Elena Ferrari. Promark: A Blockchain-Based Proximity Marketing Platform (Under Preparation).

2. **Anh-Tu Hoang**, Barbara Carminati, Elena Ferrari. Personalized Anonymization of Knowledge Graphs. TDSC (Under Review) (Q1-IF=6.4).

3. **Anh-Tu Hoang**, Barbara Carminati, Elena Ferrari. 2022. Time-Aware Anonymization of Knowledge Graphs. TOPS. (Q1-IF=2.7)

4. **Anh-Tu Hoang**, Barbara Carminati, Elena Ferrari. 2021. Privacy-Preserving Sequential Publishing of Knowledge Graphs. ICDE. (A*-AR=19.1%)

5. **Anh-Tu Hoang**, Barbara Carminati, Elena Ferrari. 2020. Cluster-based anonymization of knowledge graphs. ACNS. (B-AR=19.6%)

6. V. Nguyen, J. De Beenhouwer, S. Bazrafkan, **A-T Hoang**, S. Van Wassenbergh, and J. Sijbers. 2020. BeadNet: A Network for Automated Spherical Marker Detection in Radiographs for Geometry Calibration. CTMeeting-2020, Regensburg, Germany.

7. **Anh-Tu Hoang**, Barbara Carminati, Elena Ferrari. 2019. Cluster-Based Anonymization of Directed Graphs. CIC.

8. Hoang-Quoc Nguyen-Son, **Anh-Tu Hoang**, Minh-Triet Tran, and Isao Echizen. 2013. Anonymizing Temporal Phrases in Natural Language Text to be Posted on Social Networking Services. IWDW. (C)

9. **Anh-Tu Hoang**, Hoang-Quoc Nguyen-Son, Minh-Triet Tran, and Isao Echizen. 2013. Detecting Traitors in Re-publishing Updated Datasets. IWDW. (C)

10. **Anh-Tu Hoang**, Minh-Triet Tran, and Anh-Duc Duong. 2012. An Indexed Bottom-up Approach for Publishing Anonymized Data. CIS. (C)