

Ibrahim Sabek

CONTACT INFORMATION

32 Vassar St, Office 32-G968
 Massachusetts Institute of Technology (MIT)
 Cambridge, MA, 02139, USA

Mobile: (+1) 651-604-7295
E-mail: sabek@mit.edu
<http://people.csail.mit.edu/ibrahimsabek/>

EDUCATION

Ph.D. in Computer Science August, 2014 - January, 2020
 Computer Science and Engineering Dept., University of Minnesota, USA

- Advisor: Prof. Mohamed F. Mokbel.
- Thesis: Adopting Markov Logic Networks for Big Spatial Data and Applications.
- Research Focus: Probabilistic databases, scalable knowledge base systems, containerization systems, big spatial and spatio-temporal data management, processing and analysis.

M.Sc. in Computer Science August, 2014 - January, 2018
 Computer Science and Engineering Dept., University of Minnesota, USA

M.Sc. in Computer and Systems Engineering March, 2012 - August, 2014
 Computer and Systems Engineering Dept., Alexandria University, Egypt

- Thesis: Intelligent Hybrid Man-Machine Translation Evaluation.

B.Sc. in Computer and Systems Engineering September, 2005 - June, 2010
 Computer and Systems Engineering Dept., Alexandria University, Egypt

- Cumulative Grade: **Distinction** with Degree of Honor (GPA: 3.91/4.0).

RESEARCH EXPERIENCE

MIT, Computer Science and Artificial Intelligence Laboratory (CSAIL), MA, USA
Postdoctoral Associate **February, 2020 - Present**

- **Learned Systems Components:** I have been working on enhancing traditional data structures and algorithms through machine learning, with special focus on database management systems. This work results in papers in **CIDR'21** and **AIDB@VLDB'20**.

University of Minnesota, Computer Science and Engineering Department, MN, USA
Research Assistant **August, 2014 - January, 2020**

- **Spatial Markov Logic Networks (SMLN):** I have been working on adapting and scaling up a spatial variation of the Markov Logic Network framework to build efficient spatio-temporal analysis as well as knowledge base applications. This work results in papers in **ICDE'20**, **SIGSPATIAL Special'20**, **ACM TSAS'19** and **SIGSPATIAL'18** (has been selected as a **Best Paper Candidate**), demos in **VLDB'19** and **SIGMOD'18**, and an extended abstract in **SIGSPATIAL'19** (**Gold medal** winner of the graduate student research competition (SRC)).
- **SpatialHadoop:** I have been working on extending SpatialHadoop to support optimizing large-scale spatial queries e.g. spatial join. SpatialHadoop is an open-source MapReduce extension for Apache Hadoop with built-in support for spatial data. Recently, SpatialHadoop has been acquired under the name of **GeoJini** by eclipse foundation as one of its LocationTech projects. This work results in a paper in **SIGSPATIAL'17**, and a poster in **SIGMOD'17**.
- **Sphinx:** I have been working on supporting efficient spatial indexing and query processing inside the Apache Impala framework, under project called Sphinx. The main objective is to implement a full stack of spatial data processing, including query parser, indexer, query planner, and query executor. This work results in a paper in **SSTD'17**.

Microsoft Research, Database Group, WA, USA
Research Intern **(June, 2017 - August, 2017) (May, 2016 - August, 2016)**

- **CRA:** I have been involved in building a common run-time framework for applications that makes it easy to create and deploy distributed dataflow-style applications on top of resource managers such as Kubernetes, YARN, and stand-alone cluster execution. This project is released to the

¹Last updated at May 29, 2021.

open source community on <https://github.com/Microsoft/CRA>, and has been used to build both offline and streaming analytics platforms such as Quill and online microservice fabrics such as Ambrosia. This work results in a paper in **ICDE'19**.

- **Coral**: I have been involved in building a distributed data intensive processing layer on top of Microsoft REEF for realizing various types of computations for big data processing. These computations include seamless integration of batch, streaming and graph processing.

NEC Research Labs, Data Management Group, CA, USA

Research Intern

May, 2015 - August, 2015

- **Odyssey**: I have been involved in improving the backbone data management system of NEC, based on the Apache Hive framework. My role was embedding query optimization, and views materialization techniques inside the Odyssey stack.

Umm Al-Qura University, KACST GIS Technology Innovation Center, Mecca, Saudia Arabia

Research Assistant

January, 2014 - July, 2014

- **HajjSense and Faheem projects**: Designed, implemented and evaluated the data collection part from either sensor networks or people-centric sensing devices, e.g., GPS-equipped smartphones, or through voluntary people. Examples of collected data include network traffic data, trajectories of moving objects (e.g., vehicles), and data measurements (e.g., air quality).

Egypt-Japan University of Science and Technology, Wireless Research Group, Egypt

Research Assistant

September, 2011 - December, 2013

- **Nuzzer**: Designed, implemented and evaluated accurate multi-entity tracking solutions that use a human's effect on the Radio Frequency (RF) in WiFi environments to infer the human's presence and location. This work has been done by investigating recent wireless technologies (e.g. 802.11n) combined with solid machine learning techniques. As a result, we published papers in **IEEE TMC'14**, **IEEE WCNC'13** and **IEEE GLOBECOM'12**, a demo in **WINTECH'12**.

Vimov LLC, Alexandria, Egypt

Research and Software Engineer

January, 2013 - June, 2013

- **News Biter**: Designed, implemented and evaluated a novel hybrid Collaborative via Content recommendation technique that satisfies users needs of both news relevance and time constraints. The resulting technique has been integrated in the released iOS application "News Biter".

Microsoft Research, Advanced Technology Lab, Cairo, Egypt

External Research Collaborator

June, 2011 - June, 2012

- **CTP Project**: Designed, implemented and evaluated a platform that integrates automated evaluation of translations with human judgments to produce accurate quality estimation of large-scale translations. The resulting platform has been integrated as a **web service for Microsoft Translator Hub** that serves requests for public translation and private ones.

Microsoft Research, Microsoft Innovation Center, Cairo, Egypt

Research Assistant

October, 2010 - February, 2011

- **i2i-3D Project**: Customized the i2i technology - introduced by Microsoft Research Cambridge - on Windows Phone 7 (WP7) via efficient machine learning and pattern recognition techniques to provide privacy call mode services for Microsoft mobile platforms.

Alexandria University, Computer and Systems Engineering Department, Alexandria, Egypt

Undergraduate Research Assistant

June, 2009 - June, 2010

- **Wiki-Rec Project**: Designed, implemented and evaluated a novel approach to construct an ontology from Wikipedia categories and articles graphs to solve the problems of using traditional ontologies for the text analysis in text-based recommendation systems. Also, proposed an efficient structure for users profiles to integrate smoothly with the built ontology. As a result, we published a paper in **ISDA'10** with a **Best Paper Runner-Up Award**.

- 2021 (AIDB Workshop).
- Eleni Tzirita Zacharitou, Andreas Kipf, **Ibrahim Sabek**, Varun Pandey, Harish Doraiswamy and Volker Markl, “*The Case for Distance-Bounded Spatial Approximations*”, Conference on Innovative Data Systems Research (CIDR), Research Track, 2021.
- Varun Pandey, Alexander van Renen, Andreas Kipf, **Ibrahim Sabek**, Jialin Ding and Alfons Kemper, “*The Case for Learned Spatial Indexes*”, International Conference on Very Large Data Bases (VLDB), 2020 (AIDB Workshop).
- Ibrahim Sabek** and Mohamed F. Mokbel, “*Sya: Enabling Spatial Awareness inside Probabilistic Knowledge Base Construction*”, IEEE International Conference on Data Engineering (ICDE), Research Track, 2020.
- Ibrahim Sabek** and Mohamed F. Mokbel, “*Machine Learning Meets Big Spatial Data*”, IEEE International Conference on Data Engineering (ICDE), 2020 (Tutorial).
- Ibrahim Sabek**, Mashaal Musleh and Mohamed F. Mokbel, “*RegRocket: Scalable Multinomial Autologistic Regression with Unordered Categorical Variables Using Markov Logic Networks*”, ACM Transactions on Spatial Algorithms and Systems (TSAS), 2019.
- Ibrahim Sabek**, Badrish Chandramouli and Umar Farooq Minhas, “*CRA: Enabling Data-Intensive Applications in Containerized Environments*”, IEEE International Conference on Data Engineering (ICDE), Research Track, 2019.
- Ibrahim Sabek**, “*Flash: Scalable Spatial Probabilistic Graphical Modeling*”, ACM SIGSPATIAL Special, 2019.
- Ibrahim Sabek** and Mohamed F. Mokbel, “*Machine Learning Meets Big Spatial Data*”, International Conference on Very Large Data Bases (VLDB), 2019 (Tutorial).
- Ibrahim Sabek**, “*Adopting Markov Logic Networks for Big Spatial Data and Applications*”, International Conference on Very Large Data Bases (VLDB), 2019 (PhD Workshop).
- Ibrahim Sabek**, Mashaal Musleh and Mohamed F. Mokbel, “*Flash in Action: Scalable Spatial Data Analysis Using Markov Logic Networks*”, International Conference on Very Large Data Bases (VLDB), 2019 (Demo).
- Ibrahim Sabek**, “*Towards Scalable Spatial Probabilistic Graphical Modeling*”, ACM International Conference on Advances in Geographic Information Systems (SIGSPATIAL), 2019 (**Gold medal winner of the graduate student research competition**).
- Badrish Chandramouli, Umar Farooq Minhas and **Ibrahim Sabek**, “*CRA: A Common Runtime for Applications*”, Technical report, MSR-TR-2019-2, 2019.
- Ibrahim Sabek**, Mashaal Musleh and Mohamed F. Mokbel, “*TurboReg: A Framework for Scaling Up Spatial Logistic Regression Models*”, ACM International Conference on Advances in Geographic Information Systems (SIGSPATIAL), Research Track, 2018 (**Best paper candidate**).
- Ibrahim Sabek**, Mashaal Musleh and Mohamed F. Mokbel, “*A Demonstration of Sya: A Spatial Probabilistic Knowledge Base Construction System*”, ACM International Conference on Management of Data (SIGMOD), 2018 (Demo).
- Ibrahim Sabek** and Mohamed F. Mokbel, “*On Spatial Joins in MapReduce*”, ACM International Conference on Advances in Geographic Information Systems (SIGSPATIAL), Research Track, 2017.
- Ahmed Eldawy, **Ibrahim Sabek**, Mostafa Elganainy, Ammar Bakeer, Ahmed Abdelmotaleb and Mohamed F. Mokbel, “*Sphinx: Empowering Impala for Efficient Execution of SQL Queries on Big Spatial Data*”, International Symposium on Spatial and Temporal Databases (SSTD), Research track, 2017.
- Ibrahim Sabek**, “*Optimizing Spatial Queries in MapReduce*”, ACM International Conference on Management of Data (SIGMOD), 2017 (**Student research competition semi-finalist**).
- Ibrahim Sabek**, Moustafa Youssef and Athanasios V. Vasilakos, “*ACE: An Accurate and Efficient Multi-Entity Device-Free WLAN Localization System*”, IEEE Transactions on Mobile Computing (TMC), Volume: 14, Issue: 2, 1 February, 2015.
- Ibrahim Sabek**, Noha A. Yousri, Nagwa Elmakky and Mona Habib, “*Intelligent Hybrid Machine Translation Evaluation*”, CoRR abs/1307.1872, 2014 (**Masters thesis**).
- Heba Abdel-Nasser, Reham Samir, **Ibrahim Sabek** and Moustafa Youssef, “*MonoPHY: Mono-*

Stream-based Device-free WLAN Localization via Physical Layer Information", IEEE Wireless Communications and Networking Conference (WCNC), 2013.

Ibrahim Sabek and Moustafa Youssef, "*MonoStream: A Minimal-Hardware High Accuracy Device-free WLAN Localization System*", CoRR abs/1308.0768, 2013.

Ibrahim Sabek and Moustafa Youssef, "*Multi-entity Device-Free WLAN Localization*", IEEE Global Communications Conference (GlobeCom), 2012.

Ibrahim Sabek and Moustafa Youssef, "*SPOT Demo: Multi-entity Device-Free WLAN Localization*", The Seventh ACM International Workshop on Wireless Network Testbeds, Experimental Evaluation and Characterization (WiNTECH) in conjunction with ACM MobiCom, 2012 (Demo).

Ahmed Elgohary, Hussien Nomier, **Ibrahim Sabek**, Mohamed Samir, Moustafa Badwi and Noha A. Yousri, "*Wiki-Rec: A Semantic-based Recommender System using Wikipedia as an Ontology*". The Tenth International Conference on Intelligent Systems Design and Applications (ISDA), 2010. ²

RESEARCH FUNDING

National Science Foundation (NSF) - IIS-1907855, "*III: Small: Adopting Machine Learning Techniques for Big Spatial and Spatio-temporal Data and Applications*". PI: Mohamed Mokbel. \$500,000, 10/1/2019 - 09/30/2022 (I was a major contributor in the project design, development, and writing, but was unable to be listed as a PI/Co-PI due to graduate student status).

TEACHING EXPERIENCE

University of Minnesota, Computer Science and Engineering Department, MN, USA
Teaching Assistant **August, 2014 - May, 2018**

Duties included shared responsibility for recitations, lab tutorials, homework assignments, grading and office hours.

- CSCI-4707 Practice of Database Systems (Fall 2014, Spring 2015, Fall 2015, Spring 2018).

Alexandria University, Computer and Systems Engineering Department, Alexandria, Egypt
Teaching Assistant/Lecturer **August, 2010 - July, 2014**

Duties included shared responsibility for lectures, exams, homework assignments, grading, office hours and leading weekly computer lab exercises.

- CS-121 Programming using Java and C/C++ (Fall 2010, Fall 2012, Fall 2013).
- CS-223 Advanced Data Structures (Spring 2010, Spring 2012, Spring 2013).
- CS-221 Object-Oriented Programming (Fall 2010).
- CS-322 Database Systems (Spring 2012, Spring 2013).
- CS-E41 Special Topics on Information Systems (Fall 2010, Fall 2012, Fall 2013).
- CS-E42 Pattern Recognition and Applications (Fall 2010).
- CS-E43 Introduction to Computer Vision (Fall 2012).

EMPLOYMENT EXPERIENCE

POET-Egypt LLC, Alexandria, Egypt
Software Development Engineer **July, 2009 - September, 2009**

- Designed, developed and tested a plug-in for agile software process using scrum on Atlassian-Jira bug tracing system with JIRA API and Apache Velocity. The plug-in has been used to serve the development of the X-solutions product introduced by POET AG, Germany.

Alexandria ACM Chapter, Alexandria, Egypt
Junior Web Developer **July, 2008 - September, 2008**

- Designed, developed and tested a website to manage registered students information for the Alexandria branch of ACM. This task includes developing with different web tools such as HTML, AJAX, JavaScript and Flex.

HONORS AND AWARDS

- **UMN Best Dissertation Honorable Mention**, University of Minnesota, 2021 [Link].
- **Computing Innovation Fellowship** from Computing Research Association (CRA) and National Science Foundation (NSF), 2020 - 2022 [Link].
- **Gold Medal (First Place)** of ACM SIGSPATIAL Student Research Competition, 2019 [Link].
- **Doctoral Dissertation Fellowship** from University of Minnesota, 2019 - 2020 [Link].
- **Best Paper Candidate** in ACM SIGSPATIAL 2018, invited to TSAS special issue [Link].
- **NSF Travel Award** from VLDB 2019, ACM SIGSPATIAL 2018, 2017 and 2015.

²Authors of this paper are in alphabetical order.

- **Finalist** in ACM SIGMOD Student Research Competition, 2017.
- **ACM Student Travel Award** from ACM SIGMOD 2017, and ACM SIGSPATIAL 2019.
- **Academic Excellence Award** from University of Minnesota, Spring 2016.
- **Graduate School Fellowship** from University of Minnesota, 2014 - 2015.
- **Egypt Media Coverage** about my Masters work from Alexandria University [Link].
- **IEEE Student Travel Award** to attend IEEE WCNC 2013 and IEEE ISDA 2010.
- **Best Bachelor's Thesis Award** from Faculty of Engineering, Alexandria University, 2010.
- **Excellence Certificate of Honor** from Alexandria University, 2006 - 2010.
- **Undergraduate Research Fellowship Award** from Exchange of Students program between IAESTE Germany and Alexandria University Egypt, 2009.
- **Semi-finalist** in Microsoft Imagine Cup local competition, Egypt, 2009.

PROFESSIONAL
SERVICES AND
ACTIVITIES

- **Program Committee Member** of SIGMOD'22, EDBT'22, SIGSPATIAL'20-'21, GEOProcessing'20-'21, AutoML@ICML'20-'21, SpatialAPI@SIGSPATIAL'20.
- **Invited Review Services** for IEEE IoT'21, IEEE TKDE'20, DAPD'20, ACM TSAS'19-'21, GeoInformatica'18-'20, IEEE TMC'13-'14, IEEE TVT'12-'14, IEEE Signal Processing'13.
- **External Review Services** for SIGMOD'15-'20, VLDB'14-'20, ICDE'14-'19, SIGSPATIAL'14-'18, EDBT'17, SSTD'14-'18, MDM'16-'17, SSDBM'17.
- **Mentor** in the MIT Undergraduate Research Opportunities Program (UROP), 2020 - 2021.
- **Member** of the Grants Committee in the Council of Graduate Students (COGS), University of Minnesota, 2019.
- **Web Master** for ACM SIGSPATIAL for 2014, 2015, 2016, 2017.
- **Presenter** in the Research Showcase Exhibit, University of Minnesota, 2017.
- **Trainer** in Machine Learning summer program, Alexandria ACM Chapter, 2013.
- **Organizer** in Start Menu Day, Alexandria University, 2013.
- **Oracle CS club leader**, Alexandria University, 2012 - 2013.

REFERENCES

Available upon Requests.