Payam Siyari

Senior Data Scientist, Uber Advanced Technologies Group

Mailing Address:
Uber Advanced Technologies Group
579 20th Street
San Francisco, CA 94107
USA

Citizenship:
Iranian (U.S. Permanent Resident)

Home Page:
https://www.payamsiyari.com
https://www.linkedin.com/in/payamsiyari
https://github.com/payamsiyari

E-mail:
payamsiyari-at-gmail(dot)com

TECHNICAL TOPICS OF INTEREST

- Data Science, Machine Learning
- Maps and Geospatial Data Science
- Natural Language Processing, Grammar Induction
- Graph Mining, Network Science, Social and Information Networks
- Algorithms, Combinatorial Optimization, Algorithmic Game Theory
- Database Systems and Information Retrieval

EDUCATION

- **Georgia Institute of Technology (GATech)**,
  School of Computer Science, College of Computing, Atlanta, GA. 2014 - 2018.
  Ph.D. in Computer Science - Minor in *Statistics*.
  Advisors: Dr. Constantine Dovrolis, Dr. Bistra Dilkina
  PhD Thesis: *Optimization-Driven Emergence of Deep Hierarchies with Applications in Data Mining and Evolution*

- **Georgia Institute of Technology (GATech)**,
  School of Computer Science, College of Computing, Atlanta, GA. 2014 - 2016.
  M.Sc in Computer Science - Specialization in *Machine Learning*. **GPA: 4.00/4.00**.
  Coursework: *Machine Learning, Deep Learning for Perception, Natural Language Processing, Data and Visual Analytics, Network Science, Time Series Analysis, High Performance Computing, Computability and Algorithms, Machine Learning for Trading, Regression Analysis, Statistical Methods*

- **Sharif University of Technology (SUT)**,
  Department of Computer Engineering, Tehran, Iran. 2011 - 2013.
  M.Sc. in Computer Engineering - Software. **GPA: 19.24/20** (1st Rank).
  M.Sc. Thesis: *Network Topology Inference from Incomplete Data* (Score: **20/20**).
  Coursework: *Statistical Pattern Recognition, Data Mining, Algorithmic Game Theory, Performance Evaluation of Computer Systems, Advanced Operating Systems, Convex Optimization (Audited)*

- **Shahid Beheshti University (SBU)**,
  Department of Mathematical Sciences, Tehran, Iran. 2007 - 2011.
  B.Sc. in Computer Science. **GPA: 18.46/20** (1st Rank).
  B.Sc. Thesis: *Design and Implementation of a Native XML Database Management System* (Score: **20/20**).

PROFESSIONAL EXPERIENCES

- **Senior Data Scientist**, Uber Advanced Technologies Group - San Francisco, CA.
  Data Science Team.
  Data Scientist II, Oct. 2018 - Dec. 2019.

  - *Full-Stack Data Scientist*:
    * Deep Learning: GeoSpatial representation learning, involving CNNs on satellite image data, RNNs on temporal trip data and GNNs on road networks.
    * Data Structures and Algorithms: GeoSpatial joins and indexing, including Uber UMM, Uber H3, S2 Geometry.

- ∗ Data Engineering: Advanced SQL, Relational schema design, BigData pipeline development (Hive, Spark).
- ∗ Statistical Analysis: A/B testing, utilizing statistical testing for strategic decision making e.g. minimum amount of miles needed for deploying service.
- ∗ Data Visualization, Analytics and Dashboarding: DASH, Bokeh.
- – *@UberEngineering Blog Showcase*:
  - ∗ Power On: Accelerating Uber's Self-Driving Vehicle Development with Data
  - ∗ Searchable Ground Truth: Querying Uncommon Scenarios in Self-Driving Car Development

- **Software Engineering Intern**, Data Science Team, Uber ATG - Pittsburgh, PA.
  Supervised by Andrew Duberstein and Dr. Collin Otis, Aug. 2017 - Dec. 2017.
- **Research Intern**, Machine Learning for Document Access and Translation (ML-DAT) Team, Xerox Research Center Europe (XRCE - Now: NAVER LABS Europe) - Grenoble, France.
  Supervised by Dr. Matthias Gallé, Aug. 2015 - Dec. 2015.
- **Research Assistant**, College of Computing, Georgia Institute of Technology - Atlanta, GA.
  Supervised by Dr. Constantine Dovrolis and Dr. Bistra Dilkina, Aug. 2014 - 2018.
- **Research Assistant**, Dept. of Computer Engineering, Sharif University of Technology - Tehran, Iran.
  Under supervision of Dr. Hamid R. Rabiee, Dec. 2011 - Dec. 2013.
- **iOS Developer**, Pichak co. - Tehran, Iran, Apr. 2011 - Sep. 2011.

PUBLICATIONS
- **Journal Papers:**

  - **P. Siyari**, B. Dilkina, C. Dovrolis, **"Evolution of Hierarchical Structure and Reuse in iGEM Synthetic DNA Sequences"**, International Conference on Computational Science (ICCS'19), pp. pp 468-482, Lecture Notes in Computer Science, Vol. 11536, 2019.
  - **P. Siyari**, B. Dilkina, C. Dovrolis, **"Emergence and Evolution of Hierarchical Structure in Complex Systems"**, Dynamics On and Of Complex Networks III (DOOCN 2017) - Machine Learning and Statistical Physics Approaches, pp. 23-62, Springer Proceedings in Complexity, 2018.
  - M. Salehi, R. Sharma, M. Marzolla, D. Montesi, **P. Siyari**, and M. Magnani, **"Spreading Processes on Multilayer Networks"**, IEEE Transactions on Network Science and Engineering, Vol. 2, No. 2, pp. 65-83, 2015.
  - M. Salehi, **P. Siyari**, M. Magnani, D. Montesi, **"Multidimensional Epidemic Thresholds in Diffusion Processes over Interdependent Networks"**, Chaos, Solitons & Fractals, Vol. 72, pp. 59-67, 2015.
  - A. Fattaholmanan, H. R. Rabiee, **P. Siyari**, A. Khodadadi, and A. Soltani-Farani, **"A Peer to Peer Compressive Sensing Framework for Network Monitoring"**, IEEE Communications Letters, Vol. 19, No. 1, pp. 38-41, 2015.
  - **P. Siyari**, H. R. Rabiee, M. Salehi, and M. Eslami, **"Network Reconstruction under Compressive Sensing"**, ASE Human Journal, Vol. 1, No. 3, pp. 130-143, 2012 *(Only the top 3% papers of SocialInformatics'12 were included)*.

- **Conference Papers:**

  - **P. Siyari**, M. Gallé, **"The Generalized Smallest Grammar Problem"**, Proceedings of International Conference on Grammatical Inference (ICGI'16), PMLR 57, pp. 79-92, 2017.

- **P. Siyari**, B. Dilkina, C. Dovrolis, **"Lexis: An Optimization Framework for Discovering the Hierarchical Structure of Sequential Data"**, Proceedings of ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD'16), pp. 1185-1194, Aug. 2016 (*Oral Presentation - Acceptance Rate: 18.11%*).
- H. Mahyar, H. R. Rabiee, Z. S. Hashemifar, and **P. Siyari**, **"UCS-WN: An Unbiased Compressive Sensing Framework for Weighted Networks"**, Proceedings of Conference on Information Sciences and Systems (CISS'13), pp. 1-6, Mar. 2013.
- **P. Siyari**, H. R. Rabiee, M. Salehi, and M. Eslami, **"Network Reconstruction under Compressive Sensing"**, Proceedings of ASE/IEEE International Conference on Social Informatics, pp. 19-25, IEEE Comp. Soc., Dec. 2012.
- P. Siyari, H. Aghaeinia, and **P. Siyari**, **"Power Allocation in Relay Networks: A Stackelberg Game Approach"**, Proceedings of Majlesi Symposium on Telecommunication Devices, Feb. 2014.

- **Technical Reports:**

  - **P. Siyari**, and H. R. Rabiee, **"Fast Non-negative Matrix Factorization under KL-Divergence: A Case for Link Prediction"**, 2013.

- **Theses:**

  - **P. Siyari**, **"Optimization-Driven Emergence of Deep Hierarchies with Applications in Data Mining and Evolution"**, Computer Science PhD Thesis, School of Computer Science, College of Computing, Georgia Institute of Technology, 2018.

  - **P. Siyari**, **"Network Topology Inference from Incomplete Data"**, Computer Engineering (Software) M.Sc. Thesis, Department of Computer Engineering, Sharif University of Technology, 2013 (in Persian).

  - **P. Siyari**, **"Design and Implementation of a Native XML Database Management System"**, Computer Science B.Sc. Thesis, Faculty of Mathematical Sciences, Shahid Beheshti University, 2011 (in Persian).

HONORS AND AWARDS

- **College of Computing Travel Grant** from Graduate Student Council and Graduate Programs Office, Georgia Institute of Technology, 2018.
- **1st Rank** GPA among graduates of *Computer Engineering - Software* M.Sc. students (2011 beginners), Sharif University of Technology, 2013, Tehran, Iran.
  Remark: Because of this high rank, I was permitted to continue my studies as a PhD student without having to attend any examinations.
- **2nd Rank** according to GPA among *all* graduates of *Computer Engineering* M.Sc. students (2011 beginners), Sharif University of Technology, 2013, Tehran, Iran.
- **6th Rank** among 1,006 test takers in the Nationwide PhD Entrance Examination (Konkoor) in *Computer Engineering - Artificial Intelligence*, 2013, Tehran, Iran.
- **1st Rank Overall** among 26,195 test takers in the Nationwide M.Sc. Entrance Examination (Konkoor) in *Computer Engineering*, 2011, Tehran, Iran.
  - **1st Rank** in *"Software Engineering"*, **1st Rank** in *"Algorithms & Computations"*, **5th Rank** in *"Artificial Intelligence"*, **5th Rank** in *"Computer Architecture"*.
- **2nd Rank** among 2,132 test takers in the Nationwide M.Sc. Entrance Examination (Konkoor) in *Computer Science*, 2011, Tehran, Iran.
- **1st Rank** GPA among graduates of *Computer Science* B.Sc. students (2007 beginners), Shahid Beheshti University, 2011, Tehran, Iran.
  Remark: Because of this high rank, I was permitted to continue my studies as a M.Sc. student without having to attend any examinations.
- Awarded **Exceptional Talent Admission** for **M.Sc.** Program, Department of Computer Engineering, Sharif University of Technology, 2011, Tehran, Iran.

- Offered **Exceptional Talent Admission** for **M.Sc.** Program, Department of Mathematical Sciences, Shahid Beheshti University, 2011, Tehran, Iran.
- **34th Rank** among 1,870 test takers in the Nationwide M.Sc. Entrance Examination (Konkoor) in *Computer Science* as *a Junior Student*, 2010, Tehran, Iran.

PROJECTS

- **Software Engineering Intern**, Data Science Team, Uber Advanced Technologies Group (Uber ATG).
  Supervised by Andrew Duberstein and Dr. Collin Otis, Aug. 2017 - Dec. 2017.
  - **Self-Driving Technology Engineer (Road Analytics)**
    * Scalable automated scenario identification from autonomous driving logs.

- **Research Intern**, Machine Learning for Document Access and Translation (ML-DAT) Team, Xerox Research Center Europe (XRCE - Now: NAVER LABS Europe)[1].
  Supervised by Dr. Matthias Gallé, Aug. 2015 - Dec. 2015.
  - **The Generalized Smallest Grammar Problem:** I worked on a research problem concerning a combinatorial problem called the Smallest Grammar Problem. We extended the problem definition and proposed solutions to the situation where several sequences have to be modeled using techniques from the Minimum Description Length theory.(Main Tools: *Python, NLTK, C*)

- **Research Assistant**, College of Computing, Georgia Institute of Technology.
  Supervised by Dr. Constantine Dovrolis and Dr. Bistra Dilkina, Aug. 2014 - 2018.
  - **Lexis: An Optimization Framework for Discovering the Hierarchical Structure of Sequential Data:** We proposed a framework, referred to as Lexis, that produces an optimized hierarchical representation of a given set of target strings. The model is shown to have extensive use in optimal synthesis of DNA sequences, hierarchical structure discovery, compression and feature extraction from sequential data. (Main Tools: *Python, C*)
  - **Evo-Lexis: A Model of Optimized and Incremental Hierarchical Designs:** We are studying optimized and incremental scenarios of how hierarchical architectures, e.g. technological systems, metabolic networks, network protocol stacks, neural networks, etc. appear. To model such systems, we defined and proposed solutions for a cost minimizing hierarchical design process, involving a static and an incremental optimization problem. (Main Tools: *Python, C*)
  - **Linear-time Optimal Decomposition of DNA Sequence Data:** As the course project in CS7001 course at GATech under supervision of Prof. Constantine Dovrolis, we focused on the problem of efficient decomposition of DNA sequences by reusing the repeated parts of these sequences. We employed the longest-repeated-substring heuristic and presented a linear-time implementation using a linked suffix-array data structure. (Main Tools: *Java*)
  - **A Survey on Optimization of Submodular Functions:** As the course project in CS7001 course at GATech under supervision of Prof. Bistra Dilkina, we presented a comprehensive survey of recent results on the minimization and maximization of submodular functions.
  - **Learning Hand Features for Sign Language Recognition** *(Best Project Runner-up)***:** As the course project in CS8803-DL course at GATech, we proposed to address the spatio-temporal nature of the sign-language recognition problem by employing the approaches of 3D CNNs (with spatially pre-trained weights) and CNN-RNN models that learn features across time.

---

[1]The group is now known as LIGHT: Learning, Interacting, and Generation for Handeling Text

The main idea of the project was applying a static spatial pre-training in these architectures to learn better spatio-temporal features. (Main Tools: *Theano, NumPy*)

- **Scholarly Article Recommendation for Newbies:** As the course project in CSE6242 course at GATech, we proposed a solution to help researchers find fundamental papers when starting a new field. We developed a web-app in which we used a new centrality metric in citation networks (i.e. Path-Centrality) to do the recommendation and used an advanced UI to make the software user-friendly. (Main Tools: *Django, D3*)
- **Stock Trading Strategy Learner using Reinforcement Learning:** As the course project of CS7646 course at GATech, we implemented Q-Learning and Dyna-Q end-to-end solutions to real-world stock trading. The solution learns a trading policy and manages the trader's portfolio automatically. (Main Tools: *Python, Pandas*)

- **Research Assistant**, Dept. of Computer Engineering, Sharif University of Technology. Under supervision of Dr. Hamid R. Rabiee, Dec. 2011 - Dec. 2013.
    - **SLPMF:** We proposed a set of scalable methods for link prediction and matrix factorization on KL-divergence. The methods used ideas from dynamic programming, greedy gradient descent and parallel gradient descent and basic optimization functions were implemented in C for faster execution. (Main Tools: *MATLAB, C*)
    - **CS-NetRec:** We studied the problem of inferring the sparse diffusion network from the data resulted from the information propagation data. By formation of a linear system, I used the theory of Compressive Sensing to reconstruct the network and to explore the theoretical bounds. (Main Tools: *MATLAB, C*)
    - **Multidimensional Epidemic Thresholds:** In a collaboration with researchers from University of Bologna and Uppsala University, we proposed a new concept of multidimensional epidemic threshold characterizing diffusion processes over interdependent networks. We analytically derived and numerically illustrated the conditions for multi-layer epidemics and studied the evolution of infection density and diffusion dynamics. (Main Tools: *MATLAB, Java*)
    - **P2P-CS:** We addressed the problem of network monitoring with the assumption that all nodes require access to the entire networked data. We used the Compressive Sensing theory to derive an upper bound for the number of required measurements. (Main Tools: *MATLAB, C*)
    - **MultipleNetwork:** In a collaboration with researchers from University of Bologna and Uppsala University, MultipleNetwork is a package (written with C++) for analysis and mining of multiple(x) networks, originally developed by Matteo Magnani. In a study on multidimensional epidemic thresholds on multiplex networks, I developed the basic modules for information diffusion models. (Main Tools: *MATLAB, Java*)
    - **UCS-WN:** We studied the problem of recovering sparse link vectors with network topological constraints over weighted networks. We proposed a novel framework in the context of compressive sensing to efficiently recover sparse vectors representing the properties of the links from weighted networks. (Main Tools: *MATLAB, C*)

- **Design and Implementation of a Native XML-DBMS (B.Sc. Capstone):** Designed the static view of a Native XML-DBMS and implemented a basic DB engine. Using a version of DeweyID on XML tree, I developed a refined StAX parser for data storage & retrieval and query processing. (Main Tools: *Java*)
- **iOS Developer**, Pichak co., Apr. 2011 - Sep. 2011.
    - **VPN in Touch:** A VPN account management app (client side).

- **Haram:** An app for live broadcasting from holy shrines (client and server side).
- **Web and Graphic Design**, Caspian English Academy, Apr. 2014 - 2018.
    - A series of design projects that I do from time to time, helping my brother as the founder and manager of the institute.
    - Designed the institute's logo, banners, posters, forms, infographic projects, ID cards, certificates, website and advertisements.
    - A brief demo of my work can be found at: `http://caspian-academy.ir`.
- **Implementation of** *"Random Exam Sheet Generator from Microsoft Office Word Files"*, Sanjesh Organization, Fall 2008. (Main Tools: *C#*)

TEACHING EXPERIENCES

- **Teaching Assistant**, Georgia Institute of Technology:
    - **Computer Networks I (CS 3251 - Undergrad)**, Spring 2018.
    - **Design & Analysis of Algorithms (CS 3510 - Undergrad)**, Spring 2016.
- **Teaching Assistant**, Sharif University of Technology:
    - **Machine Learning (Grad), Social and Information Network Analysis (Grad)**, Fall 2013.
    - **Statistical Pattern Recognition (Grad), Program Development from Formal Specification (Grad)**, Spring 2013.
    - **Formal Specification and Verification (Grad)**, Fall 2012, 2013.
- **Teaching Assistant**, Shahid Beheshti University:
    - **Numerical Analysis, Probability & Statistics**, Spring 2011.
    - **Data Structures, Compilers, Theory of Computation**, Fall 2010.
    - **Advanced Programming, Numerical Linear Algebra**, Fall 2009.

SERVICES

- **External Reviewer for:**
    - IEEE Transactions on Information Forensics & Security   – PLOS ONE
        – ALENEX'17   – NetSciCom'14   – SocialCom'13   – ICEE'13,'15

SKILLS AND EXPERIENCES

Programming:	Python, Java, C, C++, Swift, Objective-C, C#
Machine Learning:	PyTorch, Tensorflow, Keras, NLTK, NeworkX, SNAP, RapidMiner
Big Data:	Hive, Spark, Amazon AWS, Pig
Database Systems:	SQLite, PostgreSQL, MySQL, Microsoft SQL Server
Web Technologies:	React JS, Django, DASH, Bokeh, D3, PHP
Scientific Computing:	MATLAB, OpenMPI, R, Octave, Mathematica
Operating Systems:	Mac OS X, Linux (Ubuntu, Debian), Windows, FreeBSD

REFERENCES Available upon request.