# TUNG HOANG

*https://www.linkedin.com/in/htt1084 • 312-730-6386 • thoang3@uic.edu*

---

**EDUCATION**

**University of Illinois at Chicago (UIC),** Chicago, IL
Doctor of Philosophy in Applied Mathematics, GPA: 3.73/4.0     *August 2017*
Master of Science in Computer Science, GPA: 3.8/4.0     *May 2017*
Master of Science in Pure Mathematics, GPA: 3.92/4.0     *May 2011*
*Selected courses*: Machine Learning, Data Mining and Text Mining, Artificial Neural Network, Artificial Intelligence I, Database Systems, Computer Systems, Operating Systems, Software Design with Java, Statistical and Probability Theory, Computer Algorithms I and II, Computational Geometry, and Graph Theory.
**Honors Program, Vietnam National University,** Hanoi, Vietnam
Bachelor of Science in Mathematics, GPA: 3.7/4.0     *May 2006*

**WORK EXPERIENCE**

**RESEARCH ASSISTANT IN BIOINFORMATICS**, UIC, Chicago, IL     *2012-present*

- Applied Artifical Neural Network (**ANN**) to detect protein-coding and non-coding regions, and to differentiate Enhancers from nonfunctional regions within DNA sequence. Designed and implemented algorithms to classify biological data using Discrete Fourier Transform (DFT), Chaos Game Representation (CGR). Written in **Matlab/ Python.**
- **PS-M:** Designed and implemented **P**ower **S**pectrum-**M**oment method to classify genomes and genes using Discrete Fourier Transform (DFT) and power spectrum. Performed cluster analysis to analyze biological sequences using **PS-M**, improved speed up to 15 times while keeping accuracy on several cases. Written in **Matlab/ R/ C++.**

**RESEARCH AND COURSE PROJECTS,** University of Illinois at Chicago     *2014-present*

- **RSA Encryption/ Decryption:** Designed and implemented fast algorithm to encrypt data using public key and decrypt data using private key. Built user-defined BigInt class to hold big integers, with methods for fast multiplication, division, greatest common divisor, modulo, exponentiation. Written in **Java**.

- **Neural Network-Digit Classification with Backpropagation:** Implemented backpropagation algorithm with gradient descent from scratch. Designed and trained a 1-hidden layer (784 x 50 x 10) neural network to classify MNIST dataset with 60000 training images and 10000 testing images. Achieved 95% accuracy on testing data. Written in **Matlab**.

- **Data Mining with Multiple Minimum Supports:** Implemented **MS-Apriori** algorithm for Frequent Itemsets Generation and Association Rule Generation. Considered the cases of multiple minimum supports, support difference constraint, and item constraints (sets of items cannot be in the same itemset; every itemset must have certain items). Written in **Java**.

- **Text Mining-Naive Bayes Classifier for Spam Detection:** Implemented a spam SMS messages classifier using Naive Bayes model, where an SMS message is considered as a bag of words. Achieved 90% accuracy. Implemented another classifier using **Logistic Regression.** Achieved 92% accuracy. Written in **Matlab.**

- **Machine Learning-Robotic Arm Dynamic Re-Training:** Designed and modified an adaptive classifier using Linear Discriminant Analysis for a single degree of freedom without affecting the accuracy or control performance for the remaining degrees of freedom in an upper limb prosthetic device. Achieved 98% accuracy. Written in **Matlab**.

- **Object Oriented Design-Monopoly Game:** Designed and implemented the board game Monopoly. Included 2 major components: (1) back end with classes to model players and different kinds of locations and functionalities on the game board and (2) front end with a model of the game with user interaction. Leveraged inheritance and polymorphism in OOP. Written in **Java**.

- **CTA L-stop Database:** Designed a relational database to store Loop's transit information from Chicago Transit Authority (CTA, http://www.transitchicago.com/data/). Displayed all the L-

stops (Loop stops), along with the stop's information such as name, lines, direction, handicap accessible, total and average ridership. Developed user interfaces using **Stored Procedures**, **Data-Binding**, **ORM**, and **LINQ**. Written in **C#/ SQL**.

- **uiTunes Database:** Designed a database with a Win Forms application that allows user to buy songs or albums. Displayed songs or albums per requested, with corresponding information such as artist(s), release date, price, number of reviews, average review. Checked if a song or an album has been purchased or not, if user has enough credit to purchase or not. Updated the database using **Transaction,** allowing rollback if errors encountered. Written in **C#/ SQL**.

- **Netflix Database**: Designed and analyzed a database containing Netflix's movies information such as: names, ids, reviews, ratings, users. Updated, deleted, inserted data into the database using **C++**. Analyzed data to extract important information such as top reviewed, top average rating, top users. Implemented another version using .NET framework: used **ADO.NET** to access the database, built GUI form of the application using **C#**. Improved the application using **N-tier** design, including Business and Data Access Tiers. Written in **C++/ C#/ SQL**.

- **Swype**: Designed and implemented an algorithm to generate a best guess with top 3 most popular alternatives from a *swype* on a QWERTY keyboard. Searched for words based on matches of the first and last characters, minimum length and appearance of all characters. Implemented another version using **F#**. Written in **C++/ C#**.

- **Convex Hulls and Segments Intersection Determination**: Implemented an algorithm to determine convex hulls of a set of points in the plane. Derived a dual algorithm to determine upper and lower envelopes of a set of lines in the plane. Built an algorithm to determine all intersection points for a set of planar segments. Written in **C/ C++**.

- **Planar Graph Determination**: Devised a naive algorithm to determine whether or not a given graph is planar using Kuratowski's theorem. Improved linear-time algorithm for planarity testing. Written in **C/ C++**.

**TEACHING ASSISTANT,** UIC, Chicago, IL                                              *2009-2015*

Leading discussions on various subjects, including Calculus, Algebra, Differential Equation, Linear Algebra, Combinatorics, Probability, and Statistics. Grading quizzes and homework, proctoring and grading midterms and finals. Tutoring at Math Learning Center on same subjects.

| | |
|---|---|
| **SKILLS** | - *Proficient:* Java, C, C++, Matlab, F#, SQL, C#, ADO.NET, Maple, Threading.<br>- *Prior experience:* Python, SAS, R, HTML5, CSS, JavaScript, Parallel Programming. |
| **PUBLICATION** | - **Hoang, Tung**, Changchuan Yin, Stephen S-T. Yau. "Numerical encoding of DNA sequences by chaos game representation with application in similarity comparison." *Genomics* (2016).<br>- **Hoang, Tung**, et al. "A new method to cluster DNA sequences using Fourier power spectrum." *Journal of theoretical biology* 372 (2015): 135-145. |

**AWARDS/ ACTIVITIES**

- Yeuk-Lam Yau-Leung Memorial Scholarship                                              *2017*
- Certificate of Peer Facilitator at the 2015 International TA Orientation at UIC        *2015*
- Student Travel Award for FNFYT, CIMPA                                          *2010, 2012*
- Coordinator of International Football Club at UIC                              *2011- present*
- Vietnam Education Foundation Scholarship                                              *2009*
- Scholarship of Rencontres du Vietnam                                                  *2006*
- National Mathematics Olympiad Gold Medal, Ministry of Education and Training, Vietnam Mathematics Society (rank 2/400)                                              *2002*