

Chad A. Williams, Ph.D.
Assistant Professor
Department of Computer Science
Central Connecticut State University

Maria Sanford Hall
1615 Stanley Street
New Britain, CT 06050
(860) 832-2719
cwilliams@ccsu.edu
<http://cs.ccsu.edu/~williams/>

RESEARCH INTERESTS

Developing machine learning techniques to address problems that require a time-sensitive approach, such as data streams. My focus is on both the theoretical aspects as well as identifying opportunities for interdisciplinary collaborations to address real-world problems.

Areas of study: machine learning, data mining, software engineering, distributed systems and transportation

EDUCATION

- Ph.D. University of Illinois at Chicago** **May 2010**
Computer Science
Thesis: *Learning Activity Patterns of Individuals*
Advisors: Peter Nelson (co-chair), Abolfazl (Kouros) Mohammadian (co-chair)
- M.S. DePaul University, Chicago, Illinois** **May 2006**
Computer Science (with distinction)
Thesis: *Profile Injection Attack Detection for Securing Collaborative Recommender Systems*
Advisor: Bamshad Mobasher
- B.S. Cornell University, Ithaca, New York** **May 1998**
Computer Science

TEACHING EXPERIENCE

Assistant Professor, Department of Computer Science **Aug. 2011 – present**
Central Connecticut State University, New Britain, CT

Tenure-track faculty member. Taught introductory and upper level computer science courses. Courses included: Intro to Computers; Intro to Internet Programming; Web Programming; Legal, Social, Ethical, and Economic Issues in Computing; and Computer Security.

Assistant Professor, Department of Mathematics and Computer Science **Aug. 2010 – June 2011**
Bemidji State University, Bemidji, MN

Tenure-track faculty member. Researched and taught introductory and upper level computer science courses, 3 each semester, including the program's capstone course. For the capstone course, initiated contact with the Boys and Girls Club of Bemidji and instructed students while they worked with the Club's director to develop a program to allow the Club to track donors per the director's specified requirements. Created syllabi for all courses, selected course textbooks, and defined a new course: Object-Oriented Software Development. Acted as the program coordinator of the computer science program and managed the department's Linux lab machines and servers. Advised 18 undergraduate students in their course selection and obtaining work after graduation. Appointed to and served on the University's Honors Council. Advisor to Computer Science Club and coached a team of students for the DigiKey Programming Competition.

Teaching Associate, Department of Computer Science **Summer 2006**
University of Illinois at Chicago, Chicago, IL

Co-instructed class of 20 students, preparing and presenting 60% of the lectures. Developed the syllabus, managed group projects, wrote and graded the final exam and directed work of the teaching assistant.

Courses Taught:

- **Computer Security**, Central Connecticut State University, 2012
- **Legal, Social, Ethical, and Economic Issues in Computing**, Central Connecticut State University, 2012
- **Introduction to Internet Programming**, Central Connecticut State University, 2011-2012
- **Web Programming**, Central Connecticut State University, Fall 2011
- **Introduction to Computers**, Central Connecticut State University, Fall 2011
- **Object-Oriented Software Development**, Bemidji State University, Spring 2011
- **Software Engineering II**, Bemidji State University, Spring 2011
- **Software Engineering I**, Bemidji State University, Fall 2010
- **Problem Solving and Computer Science**, Bemidji State University, 2010-2011
- **Linux Systems Programming**, Bemidji State University, Fall 2010
- **Co-Instructor, Introduction to Software Engineering**, Univ. of Illinois at Chicago, Summer 2006

RESEARCH EXPERIENCE

NSF IGERT Research Fellow, University of Illinois at Chicago, Chicago, IL **2006-2010**

Initiated interdisciplinary efforts aimed at applying computational methods to enable novel real-time intelligent traveler applications. Developed algorithms and techniques to allow unobtrusive learning of traveler activity and movement patterns. Led undergraduate and graduate research team in design and build of web-enabled survey and supported survey of 100 households. Spearheaded international collaborations across multiple disciplines which integrated machine learning and data mining methods with transportation planning and geographic analysis. Presented at international conferences in Ireland and Germany and delivered talks in computer science seminar series.

Research Assistant, DePaul University, Chicago, IL **2004-2006**

Research on the Secure Personalization Project explored the vulnerabilities of recommender systems and techniques for enhancing their robustness. Research included a thorough examination of the vulnerabilities of several popular collaborative recommendation algorithms and analyzed methods by which attacks could be recognized and their impact minimized. Participated in drafting a NSF project proposal. Co-authored and presented numerous papers accepted by refereed conferences and journals, including 2 best paper awards.

Project Staff, Cornell University, Ithaca, NY **1997-1998**

Research on the Ensemble Project included implementing a distributed version control system similar to CVS in order to demonstrate potential applications of a secure group communication protocol that was developed by the research group in ML, a functional language, for UNIX. Served as system administrator and wrote documentation for new users once the system became operational.

FELLOWSHIPS/AWARDS/HONORS

Dean's Research Initiative Grant **2012**

NSF Integrative Graduate Education and Research Traineeship (IGERT) Fellow **2006-2010**

Interdisciplinary graduate education spanning science, technology, engineering, mathematics and social sciences

Best Paper Award, The 8th IEEE Conference on E-Commerce Technology (CEC) 2006
Best Paper Award, DePaul Research Symposium/Midwest Software Engineering Conference 2006
Phi Kappa Phi Honor Society
Upsilon Pi Epsilon International Honor Society for the Computing and Information Disciplines

RELATED WORK EXPERIENCE

Manager, Accenture, Chicago, IL **2001-2004**
Led large teams of developers and business specialists from project planning, design and execution of project lifecycle through delivery and support of the production system. Created computational solutions for Fortune 500 companies in insurance, capital markets, banking and credit reporting. Team lead in Accenture's initial offshore development effort which involved working in India for 3 months. Responsible on several projects for coordinating numerous teams consisting of personnel from multiple consulting companies, the client, and third-party software/service providers.

Technical Architect, BlueMeteor, Inc., Chicago, IL **2000-2001**
Led development team at internet start up, developed prototypes and helped write proposals that resulted in selling projects to three Fortune 500 companies. Responsibilities included creating the training materials for all new developers hired by the company.

Consultant, Andersen Consulting (n/k/a Accenture), Chicago, IL **1998-2000**
Participated in requirements definition, application and database design, implementation, and performance testing on both Java and Microsoft web applications for multiple Fortune 500 insurance companies. Created training materials and programming standards for project teams.

JOURNAL PUBLICATIONS

"Attribute Constrained Rules for Partially Labeled Sequence Completion"

by Chad A. Williams, Peter C. Nelson and Abolfazl Mohammadian.
Advances in Data Mining - Applications and Theoretical Aspects, vol. 5633 of Lecture Notes in Computer Science, (Petra Pernert, ed.), July 2009, pp. 338 - 352.

"An Automated GPS-Based Prompted Recall Survey With Learning Algorithms"

by Joshua Auld, Chad A. Williams, Abolfazl Mohammadian and Peter C. Nelson.
Transportation Letters: The International Journal of Transportation Research, vol. 1, no. 1, Jan. 2009, pp. 59-79.

"Defending Recommender Systems: Detection of Profile Injection Attacks"

by Chad Williams, Bamshad Mobasher and Robin Burke.
Service Oriented Computing and Applications, vol. 1, no. 3, Nov. 2007, pp. 157-170.

"Toward Trustworthy Recommender Systems: An Analysis of Attack Models and Algorithm Robustness"

by Bamshad Mobasher, Robin Burke, Runa Bhaumik and Chad Williams.
ACM Transactions on Internet Technology, vol. 7, no. 4, Oct. 2007, ACM.

"Analysis and Detection of Segment-Focused Attacks Against Collaborative Recommendation"

by Bamshad Mobasher, Robin Burke, Chad Williams and Runa Bhaumik.
In *Advances in Web Mining and Web Usage Analysis*, vol. 4198 of Lecture Notes in Artificial Intelligence, (O. R. Zaiane, O. Nasraoui and P. S. Yu, eds.), 2006, pp. 96-118.

REFEREED CONFERENCE PUBLICATIONS

"Competing Hazard Model of Household Vehicle Transaction Behavior with Discrete Time Intervals and Unobserved Heterogeneity"

by Martina Z. Frignani, Joshua Auld, Abolfazl Mohammadian, Chad Williams and Peter C. Nelson.

To appear in *Proceedings of the 89th Annual Meeting of the Transportation Research Board*, (DVD), (Held at Washington, D.C.) January 2010.

“Mining Sequential Association Rules for Traveler Context Prediction”

by Chad A. Williams, Abolfazl Mohammadian, Peter C. Nelson and Sean T. Doherty.

In *Proceedings of the First International Workshop on Computational Transportation Science*, (Held at The International Conference on Mobile and Ubiquitous Systems: Networks and Services (MOBIQUITOUS 2008), Dublin, Ireland), July 2008.

“Classification Features for Attack Detection in Collaborative Recommender Systems”

by Robin Burke, Bamshad Mobasher, Chad Williams and Runa Bhaumik.

In *KDD '06: Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining*, (Philadelphia, Pennsylvania), 2006, pp. 542-547.

“The Impact of Attack Profile Classification on the Robustness of Collaborative Recommendation”

by Chad Williams, Runa Bhaumik, Robin Burke and Bamshad Mobasher.

In *Proceedings of the 2006 WebKDD Workshop*, (Held at KDD 2006, Philadelphia, Pennsylvania), Aug. 2006.

“Detection of Obfuscated Attacks in Collaborative Recommender Systems”

by Chad Williams, Bamshad Mobasher, Robin Burke, Jeff Sandvig and Runa Bhaumik.

In *Proceedings of the ECAI'06 Workshop on Recommender Systems*, (Held at the 17th European Conference on Artificial Intelligence (ECAI'06), Riva del Garda, Italy), Aug. 2006.

“Securing Collaborative Filtering Against Malicious Attacks Through Anomaly Detection”

by Runa Bhaumik, Chad Williams, Bamshad Mobasher and Robin Burke.

In *Proceedings of the 4th Workshop on Intelligent Techniques for Web Personalization (ITWP'06)*, (Held at AAAI 2006, Boston, Massachusetts), July 2006.

“Detecting Profile Injection Attacks in Collaborative Recommender Systems”

by Robin Burke, Bamshad Mobasher, Chad Williams and Runa Bhaumik.

In *Proceedings of the 8th IEEE Conference on E-Commerce Technology (CEC'06)*, (San Francisco, California), June 2006.

* *Winner of Best Paper Award*

“Evaluation of Profile Injection Attacks in Collaborative Recommender Systems”

by Chad Williams, Runa Bhaumik, Jeff Sandvig, Bamshad Mobasher and Robin Burke.

In *DePaul CTI Research Symposium / Midwest Software Engineering Conference (CTIRS/MSEC 2006)*, (Chicago, Illinois), Apr. 2006.

* *Winner of Best Paper Award*

“Segment-Based Injection Attacks Against Collaborative Filtering Recommender Systems”

by Robin Burke, Bamshad Mobasher, Runa Bhaumik and Chad Williams.

In *Proceedings of the 2005 International Conference on Data Mining (ICDM'05)*, (Houston, Texas), Nov. 2005.

“Collaborative Recommendation Vulnerability to Focused Bias Injection Attacks”

by Robin Burke, Bamshad Mobasher, Runa Bhaumik and Chad Williams.

In *Proceedings of the Workshop on Privacy and Security Aspects of Data Mining*, (Held at ICDM'05, Houston, Texas), Nov. 2005.

“Effective Attack Models for Shilling Item-Based Collaborative Filtering Systems”

by Bamshad Mobasher, Robin Burke, Runa Bhaumik and Chad Williams.

In *Proceedings of the 2005 WebKDD Workshop*, (Held at KDD 2005, Chicago, Illinois), Aug. 2005.

TALKS

“Attribute Constrained Rules for Partially Labeled Sequence Completion,” talk presented at Industrial Conference on Data Mining, Leipzig, Germany, July 2009.

“Attribute Constrained Rules: A New Approach for Missing Traveler Data,” talk presented at University of Illinois at Chicago, Department of Computer Science Colloquium, July 2009.

“Learning Travel Patterns of Individuals,” talk presented at University of Illinois at Chicago, IGERT Seminar Series, January 2009.

“Mining Sequential Association Rules for Traveler Context Prediction,” talk presented at *First International Workshop on Computational Transportation Science*, Dublin, Ireland, July 2008.

“Quickly Learning Activity and Travel Patterns of Individuals: Transfer Learning for Individual Travel Behavior Prediction,” talk presented at University of Illinois at Chicago, IGERT Seminar Series, February 2008.

“Effective Attack Models for Shilling Item-Based Collaborative Filtering Systems,” talk presented at WebKDD Workshop held at KDD 2005, Chicago, Illinois, August 2005.

OTHER PUBLICATIONS AND POSTERS

“Computational Transportation Science: An Interdisciplinary Approach to Integrating Emerging Technologies into Transportation,” by Chad A. Williams, Ouri Wolfson and Peter C. Nelson, Poster presented at 2008 NSF IGERT Project Meeting, Arlington, Virginia, May 2008.

“Genetically Evolving Optimal Neural Networks,” by Chad Williams. In *Neural Networks and Expert Systems*, The Institute of Chartered Financial Analysts of India (ICFAI), Jan. 2007.

PROFESSIONAL ACTIVITIES

President, DePaul Student Chapter of IEEE, DePaul University

Fall 2004 – Summer 2006

Member, Association for Computing Machinery (ACM)

SIGCSE, SIGKDD, SIGART, SIGMOBILE

Member, American Association for Artificial Intelligence (AAAI)

Member, IEEE

REFERENCES

Marty J. Wolf, Professor, Bemidji State University

mjwolf@bemidjistate.edu

Peter C. Nelson, Professor and Dean of Engineering, University of Illinois at Chicago

nelson@uic.edu

Abolfazl (Kouros) Mohammadian, Associate Professor, University of Illinois at Chicago

kouros@uic.edu

Bamshad Mobasher, Professor, DePaul University

mobasher@cs.depaul.edu

Robin Burke, Associate Professor, DePaul University

rburke@cs.depaul.edu

Brian Pitterle, Partner, Accenture

brian.c.pitterle@accenture.com

DISSERTATION ABSTRACT
Learning Activity Patterns of Individuals

by
Chad A. Williams
Ph.D. Candidate
Department of Computer Science
University of Illinois at Chicago

Co-Chairs:
Peter C. Nelson, Department of Computer Science and
Kouros Mohammadian, Department of Civil Engineering

Recent work has begun to examine predicting traveler patterns at an individual level through two main approaches. Transportation planners have examined an activity based analysis approach in part due to the theoretic transferability across people with similar characteristics. While these models are transferrable, the predictions are more for a type of individual than a specific individual and require a considerable amount of data from the traveler. Mobile and ubiquitous computing researchers, on the other hand, have focused on learning the patterns of individuals based on passive location observations through GPS traces of that specific individual. While the predictions of this type of model are tailored to the individual, a lengthy history of the individual's travels must be collected before significant patterns and adequate coverage emerge. This work introduces a way to combine these two approaches to learn an individual's activity patterns with limited data input required from the traveler.

This study examines how activity patterns of individuals may be quickly learned with limited traveler input. Reducing participant burden is a critical component of making this type of learning practical for applications such as an intelligent traveler's assistant (ITA). An analysis is presented of how technologies such as GPS can be combined with machine learning methods to greatly reduce the effort needed to develop an activity model of the traveler. A primary contribution of this work was developing a technique to model the activity and travel patterns of an individual on an ongoing basis that greatly reduces data collection requirements. This advance opens up new possibilities for knowing current context and projected context for applications like an ITA without overwhelming the user with data entry requirements. While this research is aimed at enhancing real-time travel applications, there are a number of other applications of this work as well. An example of this is reducing user burden in travel surveys; a design of an adaptive travel survey is introduced as one way this type of approach can be used to improve common tasks such as activity surveys. The benefits of this approach are illustrated in the reduction of participant burden in a multi-day activity based travel survey.

Due to the challenge of reducing data requirements while still developing a meaningful model of the traveler, new methods were needed. To address this goal, techniques that would allow ongoing data collection to reduce the questions asked of the traveler were examined. One way this was accomplished was through a new procedure that allowed models of the individual to be continually adapted as new data was collected. Second, a new mining algorithm was introduced that further reduced data requirements by improving predictions despite missing values. This technique proved particularly helpful in allowing non-intrusive collection methods such as GPS traces to largely replace the need for many questions as the collection period progressed. Finally, a technique was developed to use data from other surveys to improve the activity model of the individual despite differences in the city in which the other surveys were collected. The combination of all of these advances is shown to significantly reduce data requirements in an ongoing collection effort. To demonstrate the benefits of these advances, a real-world example of how this technique might be used to improve existing collection efforts was developed in the form of a multi-day activity survey. Results based on 42 households demonstrate the advantages of this work in being able to reduce the respondent burden with limited impact on the results of the survey.

In summary, this research extends current ideas and introduces algorithms and techniques for learning individual travel behavior. The focus of this research was to leverage the transferrable aspects of travel behavior and patterns to reduce learning time, while also creating a richer model of the individual traveler. This research effort introduces algorithms and techniques needed to address the problem of learning and predicting the activity needs of an individual for anticipating their associated travel demands. The results presented demonstrate that the techniques introduced in this work can make a significant impact on curtailing the impact on participants for multi-day activity learning. Advances such as these are likely to make the proposition of longer surveys or the interaction required for real-time traveler applications like an ITA much more palatable.