CSER Fall 2012 Meeting
November 4, 2012
Hilton Suites Toronto/Markham Conference Centre

CALL FOR PARTICIPATION

The Consortium for Software Engineering Research (CSER) Fall 2012 Meeting is on Sunday, November 4, 2012, the day before CASCON 2012, at the same site, the Hilton Suites hotel in Markham, Ontario.

https://www-927.ibm.com/ibm/cas/cascon/hotel.jsp

This is an open meeting to present and discuss research results in software engineering.

Preliminary Agenda

The CSER community has formed a fabulous program. This meeting features two invited speakers, 19 research talks, and a session of 19 posters at the evening reception.

Please see the preliminary agenda (pages 3 to 6) and abstracts of the talks (pages 7 to 16) and posters (pages 17 to 26).

If you wish to bring a poster to present, and it is not already listed, please contact kenw@cs.ualberta.ca.

Registration Process

Due: Tuesday, October 30, 2012.

For each principal investigator or industrial participant, please register the attending members of your research group by October 30, 2012 at the following online form:

http://etsmtl.ca/Evenements/CSER?lang=en-ca

The registration fee is $200 for faculty members and industrial participants and $75 for students.

Thanks again to François Coallier and his team at ÉTS for handling the registrations.

For accurate food planning, we will be consulting the count of registered attendees. If you have dietary restrictions, please contact mlitoiu@yorku.ca.

Please feel free to forward this document to others of your research team, your research partners, or colleagues at your institution or company who many be able to participate.

If you have any questions, please do not hesitate to contact us.

We look forward to seeing you soon in Markham!

Ken Wong & Marin Litoiu
Co-Organizers CSER Fall 2012 Meeting

kenw@cs.ualberta.ca / mlitoiu@yorku.ca

1.0 2012-10-19 kw

Agenda

08:00 – 08:30 | **Continental Breakfast**
08:30 – 08:40 | **Welcome and Introduction**

08:40 – 09:10 | **Invited Talk (New Faculty)**

    Peter Rigby

09:10 – 10:25 | **Quality, Mining, and Analysis**

T01 **Do Faster Releases Improve Software Quality? An Empirical Case Study of Mozilla Firefox**
Foutse Khomh, Tejinder Dhaliwal, Ying Zou, and Bram Adams

T02 **An Empirical Study on Factors Impacting Bug Fixing Time**
Feng Zhang, Foutse Khomh, Ying Zou, and Ahmed E. Hassan

T03 **Inferring Semantically Related Words from Software Context**
Jinqiu Yang and Lin Tan

T04 **Line Pattern Mining: A Mandatory Step in Complex Log Data Analysis**
Farhana Zulkernine, Patrick Martin, Wendy Powley, Sima Soltani, Serge Mankovskii, and Mark Addleman

•   **Session Discussion**

10:25 – 10:55 | **Nutrition Break**

10:55 – 12:30 | **Prediction, Models, and Consistency**

T05 **Software Architecture-Based Failure Prediction**
Atef Shalan and Mohammad Zulkernine

T06 **Continuous Improvement of Regulations and Compliance by Complementing Business Intelligence Tools with Goal Analysis**
Omar Badreddin

T07 **A Study of Variability Spaces in Open Source Software**
Sarah Nadi

T08 **Comprehensibility of UML-based Software Product Line Specifications – A Controlled Experiment**
Iris Reinhartz-Berger and Arnon Sturm

T09 **WSDARWIN: Web-Service Evolution as a Business Strategic Decision**
Marios Fokaefs, Eleni Stroulia, and Paul R. Messinger

•   **Session Discussion**

12:30 – 13:30 | **Lunch**

13:30 – 13:35 | **Announcements**

13:35 – 14:05 | **Invited Talk**

Smart Condo™
Eleni Stroulia

14:05 – 15:45 | **Mobile, Users, and Context**

T10 **A Large-scale Empirical Study on User Ratings of Mobile Apps**
Israel J. Mojica, Meiyappan Nagappan, Steffen Dienst, Thorsten Berger, Bram Adams, and Ahmed E. Hassan
T11 **Why Do Users Give Low Ratings to Mobile Apps?**
Hammad Khalid, Emad Shihab, Meiyappan Nagappan, and Ahmed E. Hassan
T12 **SmarterDeals: A Context-aware Deal Recommendation System based   on the SmarterContext Engine**
Sahar Ebrahimi, Norha M. Villegas, Hausi A. Müller, and Alex Thomo
T13 **Architecting Smart Applications**
Przemek Lach, Ron Desmarais, Pratik Jain, and Hausi A. Müller
T14 **Advances in Mobile Health:**
Standardization, Security, and Semantic Analysis
Kamran Sartipi and Duane Bender
• **Session Discussion**

15:45 – 16:15 | **Nutrition Break and Initial Poster Setup**

16:15 – 17:50 | **Cloud, Systems, and Security**

T15 **DCSim: A Data Centre Simulation Tool**
Michael Tighe, Gaston Keller, Michael Bauer, and Hanan Lutfiyya
T16 **An Analysis of First Fit Heuristics for the Virtual Machine Relocation Problem**
Gaston Keller, Michael Tighe, Hanan Lutfiyya, and Michael Bauer
T17 **Components with Embedded Specification-Based Security Monitors**
Muhammad Umair Khan and Mohammad Zulkernine
T18 **Surprise: User-controlled Granular Privacy and Security for Personal Data in SmarterContext**
Juan C. Muñoz, Gabriel Tamura, Norha M. Villegas, and Hausi A. Müller
T19 **Partitioning Applications for Hybrid and Federated Clouds**
Mike Smit, Mark Shtern, Bradley Simmons, and Marin Litoiu
• **Session Discussion**

17:50 – 17:55 | **Closing Remarks**

17:55 – 18:00 | **Final Poster Setup**

18:00 – 20:00 | **Poster Session, Reception, and Awards**

P01 **A Smarter Commerce Business Model based on Personal Context Information**
<u>Norha M. Villegas</u> and Hausi A. Müller

P02 **User-Controlled Privacy and Security for Personal Context Spheres in the Smart Internet**
<u>Juan C. Muñoz</u>, Gabriel Tamura, Norha M. Villegas, and Hausi A. Müller

P03 **Smart Interactive Streaming Applications**
<u>Przemek Lach</u>, Ron Desmarais, Pratik Jain, and Hausi A. Müller

P04 **Billboard — A Contextual Social Messaging System**
Przemek Lach, <u>Ron Desmarais</u>, Pratik Jain, Hausi A. Müller

P05 **Gamifying Collaborative Decision Making**
<u>Mohammad Ali Moradian</u>, Kelly A. Lyons, Maaz Nasir, and Rock Leung

P06 **Feedback Loops for Model-Based Adaptive DoS Attack Mitigation**
Cornel Barna, Mark Shtern, <u>Mike Smit</u>, Vassilios Tzerpos, and Marin Litoiu

P07 **STRATOS: A Cloud Broker**
Przemyslaw Pawluk, Bradley Simmons, <u>Mike Smit</u>, and Marin Litoiu

P08 **UML Modeling and Analysis of Power Consumption for Wireless Sensor Networks (WSNs)**
<u>John K. Jacoub</u>, Ramiro Liscano, Jeremy Bradbury, and Jared Fisher

P09 **Framework for Distributed Policy-Based Management (DBPM) In Wireless Sensor Network to Support Autonomic Behavior**
<u>Nidal Qwasmi</u> and Ramiro Liscano

P10 **Advances in Mobile Health: Standardization, Security, and Semantic Analysis**
<u>Kamran Sartipi</u> and <u>Duane Bender</u>

P11 **Managing Long-Running DBMS Queries**
<u>Mastoureh Hassannezhad</u> and Patrick Martin

P12 **A Framework for Autonomic Workload Management in DBMSs**
<u>Mingyi Zhang</u>, Patrick Martin, Wendy Powley, and Paul Bird

P13 **Querying WSDL Repositories with Grok**
<u>Douglas Martin</u>, James R. Cordy, and Thomas R. Dean

P14 **Clone Detection of JavaScript-based Malware**
<u>Saruhan A. Karademir</u>, Sylvain P. Leblanc, and Thomas R. Dean

P15 **Continuous Improvement of Regulations and Compliance by Complementing Business Intelligence Tools with Goal Analysis**
<u>Omar Badreddin</u>

P16 **Will My Patch Make It? And When?**
<u>Yujuan Jiang</u> and Bram Adams

P17 **Build System Migration in the Eclipse Ecosystem**
<u>Mathieu Bollen</u> and Bram Adams

P18 **Model Correctness Patterns - An Experience Report**
Azzam Maraee, Mira Balaban, and <u>Arnon Sturm</u>

…

Talk Abstracts

Do Faster Releases Improve Software Quality? An Empirical Case Study of Mozilla Firefox

Foutse Khomh, Tejinder Dhaliwal, and Ying Zou
(Queen's University, Canada);
Bram Adams
(École Polytechnique de Montréal, Canada)

Nowadays, many software companies are shifting from the traditional 18-month release cycle to shorter release cycles. For example, Google Chrome and Mozilla Firefox release new versions every 6 weeks. These shorter release cycles reduce the users' waiting time for a new release and offer better marketing opportunities to companies, but it is unclear if the quality of the software product improves as well, since shorter release cycles result in shorter testing periods. In this presentation, we report on our empirical study of the development process of Mozilla Firefox in 2010 and 2011, a period during which the project transitioned to a shorter release cycle. We have compared crash rates, median uptime, and the proportion of post-release bugs of the versions that had a shorter release cycle with those having a traditional release cycle, to assess the relation between release cycle length and the software quality observed by the end user.

<foutse.khomh@queensu.ca>

An Empirical Study on Factors Impacting Bug Fixing Time

Feng Zhang, Foutse Khomh, Ying Zou, and Ahmed E. Hassan
(Queen's University, Canada)

Fixing bugs is an important activity. A typical process of bug fixing consists of the following steps: 1) a user files a bug report; 2) the bug is assigned to a developer; 3) the developer fixes the bug; 4) changed code is reviewed and verified; and 5) the bug is resolved. Many studies have investigated the process of bug fixing. However, to the best of our knowledge, none has explicitly analyzed the interval between bug assignment and the time when bug fixing starts. We are blind on developer's delays when fixing bugs. We examine factors affecting bug fixing time along three dimensions: bug reports, source code involved in the fix, and code changes that are required to fix the bug. We further compare different factors by descriptive logistic regression models. Our results can help development teams better understand factors behind delays, and then improve bug fixing process.

<feng.zhang@queensu.ca>

Inferring Semantically Related Words from Software Context

Jinqiu Yang and Lin Tan
(University of Waterloo, Canada)

Code search is an integral part of software development and program comprehension. The difficulty of code search lies in the inability to guess the exact words used in the code. Therefore, it is crucial for keyword- based code search to expand queries with semantically related words, e.g., synonyms and abbreviations, to increase the search effectiveness. However, it is limited to rely on resources such as English dictionaries and WordNet to obtain semantically related words in software, because many words that are semantically related in software are not semantically related in English. This paper proposes a simple and general technique to automatically infer semantically related words in software by leveraging the context of words in comments and code. We achieve a reasonable accuracy in seven large and popular code bases written in C and Java. Our further evaluation against the state of art shows that our technique can achieve a higher precision and recall.

<j223yang@uwaterloo.ca>

Line Pattern Mining: A Mandatory Step in Complex Log Data Analysis

Farhana Zulkernine, Patrick Martin, Wendy Powley, and Sima Soltani
(Queen's University, Canada);
Serge Mankovskii and Mark Addleman
(CA Technologies Inc.)

Log files provide important information for troubleshooting complex systems. However, the structures and contents of the log data and messages vary widely. For automated processing, it is necessary to first understand the layout and characteristic of the data, which becomes very challenging when massive amount of data and messages are reported by different system components in the same log file. Existing approaches apply supervised mining techniques, consider single line messages, and return frequent patterns only. We present an approach to mine line patterns for multi-line messages using a type casting mechanism. We also mine association rules that state the contextual relationship between two successive line patterns. Our tool finds both frequent and non-frequent line patterns, words and value patterns in log files given the minimum support values. To evaluate our approach, we present a comparative study of the performances of our tool against some of the existing popular research tools.

<farhana@cs.queensu.ca>

Software Architecture-Based Failure Prediction

Atef Shalan and Mohammad Zulkernine
(Queen's University, Canada)

Depending on the role of software in everyday life, the cost of a software failure can sometimes be unaffordable. Predicting failures before their manifestation is important to assure system resilience. It helps avoid the cost of failures and enables systems to perform corrective actions prior to failure occurrences. However, effective runtime error detection and failure prediction techniques encounter a prohibitive challenge with respect to the control flow representation of large software systems with intricate control flow structures. In this talk, we present a technique for failure prediction from runtime errors of large software systems. First, aiming to avoid the possible difficulties and inaccuracies of the existing Control Flow Graph (CFG) structures, we present a Connection Dependence Graph (CDG) for control flow representation of large software systems. Second, we utilize the proposed CDG to provide a connection-based signature approach for control flow error detection. Finally, we utilize the detected errors and erroneous state parameters to predict failure occurrences and modes during system runtime. This work contributes to software reliability by providing a simple control flow representation and utilizing it to detect runtime errors and predict failure occurrences and modes with high accuracy.

<atef@cs.queensu.ca>
<mzulker@cs.queensu.ca>

Continuous Improvement of Regulations and Compliance by Complementing Business Intelligence Tools with Goal Analysis

Omar Badreddin
(University of Ottawa, Canada)

The traditional governance model focuses entirely on compliance. Processes and activities are focused on enhancing compliance levels by ensuring that regulated entities strictly follow regulations. We propose to complement the traditional model with means to allow regulatory institutions to reason about the originating regulations as well. Business Intelligence tools provide excellent insights into compliance, but do a poorer job in providing insights about prescriptive regulations that do not follow a structure that is directly suitable for database systems. We propose the use of goal modeling to provide insights about compliance levels of regulations as well. Effectively, we broaden the traditional governance model to include regulations and legislation.

<oobahy@gmail.com>

A Study of Variability Spaces in Open Source Software

Sarah Nadi
(University of Waterloo, Canada)

Many software systems are configurable allowing users to customize them according to their different needs. Supporting software variability is usually divided into three parts: the configuration space, the build space, and the code space. In this work, we explore whether the information in these spaces is consistent or not. This involves investigating how these spaces work together to ensure that variability is correctly implemented, and to avoid possible inconsistencies or anomalies. Our work identifies how variability is implemented in systems such as Linux, and initially focuses on the build system. The goals of this work include helping developers take informed decisions when designing variable software, and provide better maintenance to already existing configurable systems.

<snadi@uwaterloo.ca>

Comprehensibility of UML-based Software Product Line Specifications – A Controlled Experiment

Iris Reinhartz-Berger
(University of Haifa, Israel);
Arnon Sturm
(Ben-Gurion University of the Negev, Israel)

Software Product Line Engineering deals with developing artifacts that capture the common and variable aspects of software product families. Domain models, which are one kind of such artifacts, need to specify commonality and variability and guide the reuse of the artifacts in particular software products. Although different modeling methods have been proposed to support these capabilities, the assessment of these methods is still in an inceptive stage. In this work, we examined the comprehensibility of domain models. In particular, we conducted a controlled experiment in which 116 undergraduate students were required to answer comprehension questions regarding a domain model that was equipped with explicit reuse guidance and/or variability specification. We found that explicit specification of reuse guidance within the domain model helped understand the model, whereas explicit specification of variability increased comprehensibility only to a limited extent. Explicit specification of both reuse guidance and variability often provided intermediate results. All these results were perceived in different UML diagram types and different aspects.

<sturm@bgu.ac.il>

WSDARWIN: Web-Service Evolution as a Business Strategic Decision

Marios Fokaefs, Eleni Stroulia, and Paul R. Messinger
(University of Alberta, Canada)

Web Services, and the systems based on them, are fundamentally distributed in nature. Furthermore, service-oriented systems are composed through their interfaces, hiding service-implementation details from clients. These facts hinder the evolution of such systems. Therefore, it becomes the responsibility of the provider to take into account the impact that a service change may have on client applications, in addition to considering the potential benefits they can potentially gain from the evolution of the service. In this presentation, we demonstrate a theoretic framework to support the provider's decisions in the evolution process. We employ Game Theory to construct a decision tree, which will help service providers not only make the best decisions concerning the evolution of services but also understand why these decisions are desirable.

<fokaefs@ualberta.ca>

A Large-scale Empirical Study on User Ratings of Mobile Apps

Israel J. Mojica and Meiyappan Nagappan
(Queen's University, Canada);
Steffen Dienst and Thorsten Berger
(University of Leipzig, Germany);
Bram Adams
(École Polytechnique de Montréal, Canada);
Ahmed E. Hassan
(Queen's University, Canada)

To help users decide which apps to install, app stores like Google Play leverage user ratings. Such ratings have a high correlation with download counts, yet little is known about their dynamics: What are common rating levels? Can an app recover from bad ratings? Since a good understanding of such dynamics is essential for developers to ensure the success of their apps, we mined the user ratings of 238,198 versions of 128,195 free mobile apps in the Google Play store in 2011. We found that most apps have a high median global-rating (across all versions) of 4.1 out of 5. Surprisingly, the global-rating is resilient to fluctuations once an app has gathered a substantial number of raters, i.e., developers can safely experiment with new ideas. We also studied the apps' local-rating (i.e., for a specific version), and built accurate models to predict an increase in local-rating for a new app version.

<adams.bram@gmail.com>

Why Do Users Give Low Ratings to Mobile Apps?

Hammad Khalid
(Queen's University, Canada);
Emad Shihab
(Rochester Institute of Technology, USA);
Meiyappan Nagappan and Ahmed E. Hassan
(Queen's University, Canada)

The quality of mobile apps is becoming an increasingly important issue. However, much of the research on quality of mobile apps (and even regular apps) is done from the perspective of developers, without considering the user's perspective. Therefore, in this paper we explore why users give low ratings to mobile apps by qualitatively studying a statistically representative sample of more than 7,400 low-rating user-comments for 24 mobile apps. Our manual analysis uncovers 12 types of user complaints about mobile apps. Functional errors, requests for additional features, and app crashes are examples of the most common complaints. We also find that the most common complaints are not the most negatively-impacting ones (i.e., the complaints associated with the lowest ratings). Privacy or ethical issues, and hidden app costs are two of the complaints with the most negative impact on the overall rating of an app. We find that users attributed their complaint to updating the app in almost 11\% of the studied reviews -- highlighting the importance of regression testing in mobile development. We also find that the types of complaints are mostly the same across different mobile platforms (i.e., iOS and Android). Our study provides a unique perspective on the evaluation of mobile app quality by users, which can help practitioners build better quality apps.

<emad.shihab@rit.edu>

SmarterDeals: A Context-aware Deal Recommendation System based on the SmarterContext Engine

Sahar Ebrahimi, Norha M. Villegas, Hausi A. Müller, and Alex Thomo
(University of Victoria, Canada)

Daily-deal applications are popular implementations of on-line advertising strategies that offer products and services to users based on their personal profiles. The current implementations are effective but can frustrate users with irrelevant deals due to stale profiles. To exploit these applications fully, deals must become smarter and context- aware. This talk presents SmarterDeals, our deal recommendation system that exploits users' changing personal context information to deliver highly relevant offers. SmarterDeals relies on recommendation algorithms based on collaborative filtering (CF), and SmarterContext, our adaptive context management framework. SmarterContext provides SmarterDeals with up-to-date information about users' locations and product preferences gathered from their past and present web interactions. The validation results demonstrate the suitability of our approach. For many deal categories the accuracy of SmarterDeals is between 3% and 8% better than the approaches used as baselines. For some categories, and in terms of multiplicative relative performance, SmarterDeals outperforms by 173.4%, and 37.5% on average.

<nvillega@cs.uvic.ca>

Architecting Smart Applications

Przemek Lach, Ron Desmarais, Pratik Jain, and Hausi A. Müller
(University of Victoria, Canada)

Consumers expect a high degree of personalization in their applications. To deliver on these expectations developers have to build applications that are smart. Smart applications need to deliver a personalized experience that can scale to millions at runtime. These requirements present challenges for how we architect our systems. We propose one such architecture for personalized video streaming that aims to address some of these requirements.

<przemek@przemeklach.com>

Advances in Mobile Health: Standardization, Security, and Semantic Analysis

Kamran Sartipi
(University of Ontario Institute of Technology, Canada)
Duane Bender
(Mohawk College, Canada)

We present an overview of requirements for the new generation of mobile health technology as a fast-growing application domain with major impact on health and medical services for rural regions and homecare patients. This technology takes advantage of fast and sophisticated smart mobile devices, variety of cheap and wireless body sensors, more secure communication channels, easy access to large medical records, and robust decision making algorithms. Recently, the roadblock in integration of medical information systems, caused by complexity of HL7-v3 standards, is being removed by using a new RESTful approach. We briefly present the current research problems in this field, including real-time security of information communication, agent based and context-aware semantic analysis, and communication of extra-large diagnostic imaging. Our current research agenda will provide an experimental cloud infrastructure for mobile and service communication, located at Mohawk College MEDIC lab, with a close collaboration with Smart Software Systems Lab at UOIT.

<Kamran.Sartipi@uoit.ca>
<duane.bender@mohawkcollege.ca>

DCSim: A Data Centre Simulation Tool

Michael Tighe, Gaston Keller, Michael Bauer, and Hanan Lutfiyya
(University of Western Ontario, Canada)

Computing today is shifting from hosting services in servers owned by individual organizations to data centres providing resources to a number of organizations on a shared infrastructure. Managing such a data centre presents a unique set of goals and challenges. Through the use of virtualization, multiple users can run isolated virtual machines (VMs) on a single physical host, enabling higher server utilization, reduced power consumption, and reduced infrastructure costs. To meet constantly changing workload levels, resources need to be dynamically reallocated to VMs. Algorithms to perform dynamic reallocation are open research problems. Experimenting with such algorithms on the data centre scale is impractical. Thus, there is a need for simulation tools to allow rapid development and evaluation of data centre management techniques. We present DCSim, an extensible simulation framework for simulating a data centre hosting an Infrastructure as a Service cloud.

<mtighe2@csd.uwo.ca>

An Analysis of First Fit Heuristics for the Virtual Machine Relocation Problem

Gaston Keller, Michael Tighe, Hanan Lutfiyya, and Michael Bauer
(University of Western Ontario, Canada)

In recent years, data centres have come to achieve higher utilization of their infrastructure through the use of virtualization and server consolidation (running multiple application servers simultaneously in one physical server). One problem that arises in these consolidated environments is how to deal with stress situations, that is, when the combined demand of the hosted virtual machines (VMs) exceeds the resource capacity of the host. The VM Relocation problem consists of determining which VMs to migrate and to which hosts to migrate them, so as to relieve the stress situations. In this work, we propose that the order in which VMs and hosts are considered for migration results in better outcomes, depending on the situation and the data centre's business goals. We evaluate and compare a set of First Fit-based relocation policies, which consider VMs and hosts in different order. We present simulation results showing that the policies succeed to different extents depending on the scenario and the metrics observed.

<gkeller2@csd.uwo.ca>

Components with Embedded Specification-Based Security Monitors

Muhammad Umair Khan and Mohammad Zulkernine
(Queen's University, Canada)

Security specifications have to be correctly implemented to avoid vulnerabilities in the developed software. However, some security vulnerabilities may still be present when the software is deployed. Security vulnerabilities in a software component can compound the damage as the component will be reused multiple times in different environments. Given that, a software component's behavior must be continuously observed against its security specifications after deployment to identify incorrect implementations. To overcome the aforementioned issues, we propose that a software component should have an embedded specification-based security monitor. We first identify the activities necessary to develop a software component with an embedded specification-based security monitor. We also elaborate upon the design (constituent parts) and the working (interactions between the constituent parts and the component) of the embedded monitor. Finally, we propose security monitoring techniques to observe different constraints and behaviors imposed by security specifications. We evaluate our proposed approach by embedding security monitors in component-based software (CBS). These CBS are compared with their counterparts that have external monitors with respect to design complexity, size, and performance. The viability of the proposed security monitoring techniques is assessed by calculating the size, memory usage, and monitoring overhead.

<umair@cs.queensu.ca>

<mzulker@cs.queensu.ca>

Surprise: User-controlled Granular Privacy and Security for Personal Data in SmarterContext

Juan C. Muñoz and Gabriel Tamura
(Icesi University Cali, Valle del Cauca, Colombia);
Norha M. Villegas and Hausi A. Müller
(University of Victoria, Canada)

The Smart Internet relies on the exploitation of information obtained from interactions of users with web applications. A critical aspect for its success is the adoption of mechanisms that guarantee the protection of information sensitive to users. In this talk we present Surprise, our solution to empower users with privacy and data security control for the access to their information, stored in Personal Context Sphere repositories. These repositories are defined and maintained by SmarterContext, our Smart Internet infrastructure that improves the quality of user experience in their interactions with web applications. Surprise (i) allows users to configure access permissions to their sensitive personal information to third parties, selectively and with different levels of granularity; (ii) supports changes in these configurations at runtime to add or remove third parties or permissions, and (iii) realizes partial encryption to share non-sensitive data with not explicitly authorized third parties, while protecting user identity.

<tamura.gabriel@gmail.com>

Partitioning Applications for Hybrid and Federated Clouds

Mike Smit, Mark Shtern, Bradley Simmons, and Marin Litoiu
(York University, Canada)

On-demand access to computing resources as-a-service has the potential to allow enterprises to temporarily scale out of their private data center into the infrastructure of a public cloud provider during times of peak demand. However, concerns about privacy and security may limit the adoption of this technique. We describe an approach to partitioning a software application (particularly a client-facing web application) into components that can be run in the public cloud and components that should remain in the private data center. Static code analysis is used to automatically establish a partitioning based on low-effort input from the developer. Public and private versions of the application are created and deployed; at run-time, user navigation proceeds seamlessly with requests routed to the public or private data center as appropriate. We present implementations for both Java and PHP web applications, tested on sample applications.

<msmit@cs.ualberta.ca>

Poster Abstracts

A Smarter Commerce Business Model based on Personal Context Information

Norha M. Villegas and Hausi A. Müller
(University of Victoria, Canada)

User interactions in on-line shopping are valuable sources of context information useful to enhance shopping experiences. SmarterContext, a novel context manager that discovers meaningful context from user web interactions, has great potential to develop new business models that benefit buyers, retailers and cloud infrastructure providers in the smarter commerce realm.

<nvillega@cs.uvic.ca>

User-Controlled Privacy and Security for Personal Context Spheres in the Smart Internet

Juan C. Muñoz and Gabriel Tamura
(Icesi University Cali, Valle del Cauca, Colombia);
Norha M. Villegas and Hausi A. Müller
(University of Victoria, Canada)

A critical aspect for the Smart Internet success is the adoption of mechanisms to protect users' sensitive information in web interactions. Surprise is our solution to empower users with privacy and data security control to access, transport and store their sensitive information, in personal context repositories, through the SmarterContext infrastructure.

Smart Interactive Streaming Applications

Przemek Lach, Ron Desmarais, Pratik Jain, and Hausi A. Müller
(University of Victoria, Canada)

Today consumers expect a high degree of customization of streamed media. One such innovative technology is interactive streaming. These applications posit formidable engineering challenges. We showcase innovative smart applications to manage video streams using situational context. Our implementation is based on OpenFlow, HTML5, QR codes, video analysis, and mobile platforms.

<przemek@przemeklach.com>

Billboard — A Contextual Social Messaging System

Przemek Lach, Ron Desmarais, Pratik Jain, Hausi A. Müller
(University of Victoria, Canada)

Billboard is a public social messaging system. Users can view and publish information based on their location to a publicly viewable billboard. The billboards are contextual in that they display local information on the environment. Environmental information includes the users who are near or have an interest in the billboard along with local information such as current events relevant to the geographical area and localized advertisements.

<ron.desmarais@gmail.com>

Gamifying Collaborative Decision Making

Mohammad Ali Moradian, Kelly A. Lyons, and Maaz Nasir
(University of Toronto, Canada);
Rock Leung
(SAP)

Engaging people to participate fully in collaborative decision-making activity in a software application can be challenging. People are very busy, juggling competing demands for their time. Gamification has been used in a variety of environments to incent participation and increase participation (Thom, 2012). In this project, we investigate how gamification can be used to incent and motivate people to participate in and contribute to collaborative decision-making activities.

<moradian.ali@gmail.com>

Feedback Loops for Model-Based Adaptive DoS Attack Mitigation

Cornel Barna, Mark Shtern, <u>Mike Smit</u>, Vassilios Tzerpos, and Marin Litoiu
(York University, Canada)

Denial of Service (DoS) attacks overwhelm online services, preventing legitimate users from accessing a service, often with impact on revenue or consumer trust. Approaches exist to filter network-level attacks, but application-level attacks are harder to detect at the firewall. Filtering at this level can be computationally expensive and difficult to scale, while still producing false positives that block legitimate users. We present a model-based adaptive architecture and algorithm for detecting DoS attacks at the web application level and mitigating them. Using a performance model to predict the impact of arriving requests, a decision engine adaptively generates rules for filtering traffic and sending suspicious traffic for further review, where the end user is given the opportunity to demonstrate they are a legitimate user. If no legitimate user responds to the challenge, the request is dropped. Experiments performed on a scalable implementation demonstrate effective mitigation of attacks launched using a real-world DoS attack tool.

<<msmit@cs.ualberta.ca>>

STRATOS: A Cloud Broker

Przemyslaw Pawluk, Bradley Simmons, <u>Mike Smit</u>, and Marin Litoiu
(York University, Canada)

This poster introduces a cloud broker service (STRATOS) which facilitates the deployment and runtime management of cloud application topologies using cloud elements/services sourced on the fly from multiple providers, based on requirements specified in higher level objectives. Its implementation and use is evaluated in a set of experiments.

<<msmit@cs.ualberta.ca>>

UML Modeling and Analysis of Power Consumption for Wireless Sensor Networks (WSNs)

John K. Jacoub, Ramiro Liscano, Jeremy Bradbury, and Jared Fisher
(University of Ontario Institute of Technology, Canada)

Wireless Sensor Networks (WSNs) systems are deployed to monitor specific phenomena. The design of WSNs is prone to errors and debugging and is very challenging due to the complex interactions of software components in a sensor node. Moreover, WSNs systems have limited power sources which lead to the necessity of minimizing power consumption utilization during the design. This poster presents a set of software patterns that can be used as a basis for software design of a WSN. The UML is used to capture the hardware and the software components of a WSN system and this in turn is used for power consumption analysis of the WSN during the early stages of the development cycle. The WSN modelling patterns are justified by applying them to two types of WSN systems, a typical multi-hop field deployment and another non-typical WSN-RFID WSN that integrates RFID with sensor nodes in order to support authenticated point to point communication with a sensor node.

<John.Khalil@uoit.ca>

Framework for Distributed Policy-Based Management (DBPM) In Wireless Sensor Network to Support Autonomic Behavior

Nidal Qwasmi and Ramiro Liscano
(University of Ontario Institute of Technology, Canada)

Wireless sensor networks (WSN) usually work in a heterogeneous environment, which makes a sensor's node very difficult to detect, access and manage. Therefore, there is a need for autonomic behavior to overcome these environmental challenges. A general way of implementing autonomic behavior in distributed systems is through the use of policies. However the conventional policy frameworks are generally too heavy to execute in the sensor node. Thus the goal of our research is to create a framework for distributed policy-based management in WSNs. Our proposed framework is expected to extend the WSN management functionalities compared with conventional policy management system like Finger/Finger2; it also conceals the complexity of administrating policies operations from the users by streamline the processes; finally, it overcomes the flaw in the existing frameworks about policy execution orders in some cases where multi-policies are required to ensure consistency and persistence.

<Nidal.Qwasmi@uoit.ca>

Advances in Mobile Health: Standardization, Security, and Semantic Analysis

Kamran Sartipi
(University of Ontario Institute of Technology, Canada)
Duane Bender
(Mohawk College, Canada)

We present an overview of requirements for the new generation of mobile health technology as a fast-growing application domain with major impact on health and medical services for rural regions and homecare patients. This technology takes advantage of fast and sophisticated smart mobile devices, variety of cheap and wireless body sensors, more secure communication channels, easy access to large medical records, and robust decision making algorithms. Recently, the roadblock in integration of medical information systems, caused by complexity of HL7-v3 standards, is being removed by using a new RESTful approach. We briefly present the current research problems in this field, including real-time security of information communication, agent based and context-aware semantic analysis, and communication of extra-large diagnostic imaging. Our current research agenda will provide an experimental cloud infrastructure for mobile and service communication, located at Mohawk College MEDIC lab, with a close collaboration with Smart Software Systems Lab at UOIT.

<Kamran.Sartipi@uoit.ca>
<duane.bender@mohawkcollege.ca>

Managing Long-Running DBMS Queries

Mastoureh Hassannezhad and Patrick Martin
(Queen's University, Canada)

Long-running, complex queries, such as those found in Business Intelligence (BI) workloads, can have a negative impact on the performance of a database system. Query unpredictability can result from data skew, poorly-written SQL queries, badly optimized plans and even resource contention. Recognizing errant long-running queries and taking appropriate action can minimize the effects on the workload. We investigate the use of a Query Progress Indicator (PI) to assist in determining how best to handle the management of currently executing queries.

<mh@cs.queensu.ca>

A Framework for Autonomic Workload Management in DBMSs

Mingyi Zhang, Patrick Martin, and Wendy Powley
(Queen's University, Canada);
Paul Bird
(IBM Canada)

In today's data server environments, multiple types of workloads can be mixed and present in a system simultaneously. Workloads may have different levels of business importance and unique performance objectives. An autonomic workload management system dynamically controls the flow of the workloads to help the database management systems (DBMSs) achieve the performance objectives. In this poster, we present a framework and a prototype implementation for autonomic workload management in DBMSs. The framework provides the ability to meet performance objectives of workloads with multi-level business importance, and to protect database systems against performance failure in a workload mix data server environment. The prototype system is implemented on top of IBM® DB2® Workload Manager. A set of experiments is conducted using the prototype implementation on DB2 databases to illustrate the effectiveness of our approaches.

<myzhang@cs.queensu.ca>

Querying WSDL Repositories with Grok

Douglas Martin, James R. Cordy, and Thomas R. Dean
(Queen's University, Canada)

In this work, we present an approach for querying a WSDL (Web Service Description Language) repository using Grok, an engine with its own unique language that makes it possible to query a factbase of binary relations. Using TXL, we extract a set of facts from a WSDL repository, and use Grok to ask common questions about the web services in the repository.

<douglashmartin@gmail.com>

Clone Detection of JavaScript-based Malware

Saruhan A. Karademir
(Queen's University, Canada);
Sylvain P. Leblanc
(Royal Military College of Canada, Canada);
Thomas R. Dean
(Queen's University, Canada)

JavaScript-borne malware is a primary vector of attack in the modern web. Currently, there are signature-based detection mechanisms that can provide a defense against these attacks. However, these systems are not particularly flexible and small changes to the malware can hide them from these tools. Some malware is also obfuscated by in-source compression engines that extract the actual malicious code during execution-time. Our research addresses this challenge by building a PDF-based JavaScript malware detection pipeline using the NiCad Code Clone Detection tool. This clone-detection system provides increased flexibility by taking advantage of JavaScript's inherent availability of source. Compression-based malware is dynamically detected and extracted for further detection using the V8 JavaScript execution engine. We discuss the flexibility and performance questions that rise from using clone-detection in lieu of traditional detection mechanisms.

<7sak@queensu.ca>

Continuous Improvement of Regulations and Compliance by Complementing Business Intelligence Tools with Goal Analysis

Omar Badreddin
(University of Ottawa, Canada)

The traditional governance model focuses entirely on compliance. Processes and activities are focused on enhancing compliance levels by ensuring that regulated entities strictly follow regulations. We propose to complement the traditional model with means to allow regulatory institutions to reason about the originating regulations as well. Business Intelligence tools provide excellent insights into compliance, but do a poorer job in providing insights about prescriptive regulations that do not follow a structure that is directly suitable for database systems. We propose the use of Goal modeling to provide insights about compliance levels of Regulations as well. Effectively, we broaden the traditional governance model to include regulations and legislation.

<oobahy@gmail.com>

Will My Patch Make It? And When?

Yujuan Jiang and Bram Adams
(École Polytechnique de Montréal, Canada)

The Linux kernel is maintained by thousands of volunteers who actively submit patches in the hope that their new feature or bug fix will make it into the next kernel release. However, not every patch makes it and of those that do, some patches require a lot more reviewing and testing than others. In order to support the volunteers in understanding which patches are worthwhile to pursue, we build models of the probability that a certain patch submission will be accepted and of the time it will take to get in. Our approach traces back patches all the way from their accepted version to the initial emails discussing them, and collects attributes from the emails and patches. Then, we build classification models to explore the relationship between the collected attributes and the outcome of the patch. We built and evaluated models on data of 4 years of kernel development.

<yujuan.jiang@polymtl.ca>

Build System Migration in the Eclipse Ecosystem

Mathieu Bollen and Bram Adams
(École Polytechnique de Montréal, Canada)

As the source code of a project evolves, the build system responsible for compiling it has to evolve as well, which can lead to a build system so complex that it takes too much time to maintain and starts to malfunction. As a seemingly "easy" way out, projects like KDE and the Linux kernel tend to migrate their build system towards more powerful build system technologies. However, such migrations involve great risks, challenges and costs, which are still largely undocumented. In order to uncover these risks, we are currently analyzing how sub-projects of the Eclipse project migrated recently toward Eclipse's new build system technology, Tycho. As a first step, we are comparing size and complexity measures before and after migration to learn whether or not the migration of the build system technology really improved the build system's design.

<mathieu.bollen@gmail.com>

Model Correctness Patterns - An Experience Report

Azzam Maraee, Mira Balaban, and Arnon Sturm
(Ben-Gurion University of the Negev, Israel)

Models are the backbone of the emerging Model Driven Engineering approach, whose major theme is development of software via repeated model transformations. The quality of models used in such a process affects not only the final result, but also the development process itself. In order to achieve high quality models, developers must be equipped with the awareness to model design problems and the ability to identify and correct such models. In this work we observe the role of class diagram correctness patterns as an instrument for introducing awareness to modeling problems and for improving class diagram modeling. To support that notion a catalog of correctness and quality design (anti)-patterns for class diagrams was developed. The patterns in that catalog characterize problems, analyze their causes, and provide repairing advice. In addition, an experiment of using the anti-patterns for identifying correctness and quality problems in class diagram is described.

<sturm@bgu.ac.il>

A Methodology for Integrating Security Policies within the Software Development Process

Jenny Abramov, Omer Anson, Michal Dahan, Peretz Shoval, and Arnon Sturm
(Ben-Gurion University of the Negev, Israel)

Security in general and database in particular, are crucial for organizations. While functional requirements are defined in the early stages of the development process, non-functional requirements such as security tend to be neglected or dealt with only at the end of the development process. Various efforts have been made to address this problem; however, none of them provide a complete framework to guide, enforce and verify the correct design of security policies, and eventually generate code from that design. In this work we present a methodology that addresses these gaps and assists developers to design and implement secure databases that comply with the organizational security policies. The methodology is supported by a CASE tool. The use of the proposed methodology was evaluated in an empirical experiment and two case studies.

<sturm@bgu.ac.il>