

ABSTRACTS-YEAR 2003

THESIS

COMPUTER SCIENCE

CS-03-01

WEB SERVICES WITH BEHAVIOR SPECIFICATIONS

Mr. Janeson Chakajnarodom

Prof. Vilas Wuwongse

As the number of services is always increasing, dynamic service composition systems can take advantage of the availability of a wide variety of services. A current trend is to model the structure and express the logic of a composite web service using a business process modeling language tailored for web services. A flow language is used as a representation of Web Services behaviors, and as the service implementation. The behavioral representation of a Web Service is an important task of a flow language. It provides the information that enables the users to effectively use the service and understand the control flow and logic underlying the service. This study suggests Web Services with Behavior Specification (WSBS) – a declarative approach to the construction of behavioral specifications by means of a unified modeling language XML Declarative Description (XDD) and an XML-based declarative programming language XML Equivalent Transformation (XET). By employment of XDD's language expressiveness and inference capability as well as XET's computation and query- and document- processing, WSBS introduces a new paradigm for composing a new Web service from existing Web services according to user's specifications from given a set of Web services and descriptions of some tasks or goals to be achieved with more specific requirements and behavior specifications.

CS-03-02

A CASE STUDY IN NETWORK MANAGEMENT USING DIRECTORY-ENABLED NETWORK

Ms. Punyaporn Prangjarat

Prof. Phan Minh Dung

The goals of network management applications are to monitor and to report the status of network devices in order to reduce the tasks of administrator. Most current of network management solutions are not based on information model and do not have any kind of network repository. Thus, when the network becomes large it will be more complex and difficult to manage.

This thesis proposed and implemented a new mechanism for developing network management application that applies Directory Enabled Network (DEN) to existing network management application along with Java technologies. It is powerful enough for network administrator to fix problems that arise from hardware level or system level of network devices.

For the stages of this study, we proposed the analysis, design, and implementation using the Object-Oriented methodology based on principles and functions of DEN and network management system. Initially, the problems of system are analyzed for designing a framework that combines the DEN model and network management model to perform a new paradigm of network management application. Then analysis and design of the CSIM network management system, a case study of this thesis, are discussed. Finally, the case study of this thesis is implemented for CSIM network management under Microsoft Windows platform.

CS-03-03

WEB SERVICE MODELING LANGUAGE

Ms. Xiu Feng Lu

Prof. Vilas Wuwongse

Developing a good and high quality modeling language for Web Service systems is as essential as having a blueprint for a large building. However, current existing modeling languages are limited in their expressive power, and hence cannot describe intelligent and dynamic Semantic Web Services. They also lack the ability to specify automated Web Services compositions.

In this study, a new XML-based modeling language for the description of Web Services and their compositions—Web Service Modeling Language (WSML) is designed. WSML aims to make Web services computer-interpretable, which can be described with sufficient information to enable 'automated' Web Service composition and execution. First, μ -Calculus theory is employed to provide necessary principles and standards for the design guideline. Secondly, two popular modeling languages—DAML-S and WSFL are used as the design prototype. Finally, XDD rules embedded in WSML are proposed to supply stronger interoperability.

At the end of the design work, WSML is evaluated by comparing with its similar languages (DAML-S and WSFL). The results reveal that WSML is a very straightforward and effective language to describe process models.

CS-03-04

FINDING ISOMORPHISM BETWEEN RDF GRAPHS

Mr. Sithu Naing

Prof. Ramakoti Sadananda

Resource Description Framework (RDF) recommended by the World Wide Web Consortium, is to model meta-data about the resources of the web. RDF is a fundamental lower layer of the semantic web. It describes graphs of statements about resources. Within RDF graphs, anonymous resources are considered as unlabelled in a graph. In this thesis, find out the equality of RDF graphs and traditional graphs. And show that Interactive vertex classification algorithm which is a standard graph isomorphism algorithm can be applied to check the isomorphic between RDF graphs effectively.

CS-03-05

SELF-ORGANIZING MAPS FOR BROWSING FAQs

Mr. Le Thanh Liem

Prof. Ramakoti Sadananda

This research is on the automation of Frequently Ask Questions (FAQs) faced by a telephone enquiring system. This thesis considered the Vietnam Telephone Call Center known as 1080 Net Enquiry System. This thesis explored the use of Self-Organizing Maps (SOMs) to group the FAQs to identify appropriate responses.

CS-03-06

A CASE STUDY IN TRUST MANAGEMENT: A PROPOSAL FOR A BETTER BUSINESS BUREAU (BBB) OFFICE IN HO CHI MINH CITY

Mr. To Dinh Vinh

Prof. Phan Minh Dung

Trust Management (TM) is a promising approach for authorization and access control in distributed system, based on signed distributed policy statements expressed in policy language. Role-based Trust management (RT), a recently proposed framework is claimed to have combined strengths of previous TM systems. The RT framework is a family of Role-based Trust management languages for representing policies and credentials in distributed authorization.

In this thesis, we study the concept and basic principles of TM, summarize existing Trust management systems and examine languages of the RT framework. We also analysis and propose the membership standards of Better Business Bureau (BBB) of HCMC, Vietnam, as a case study of Trust management system. Finally, this case study is represented in the RT framework.

CS-03-07

LOGIC PROGRAMMING-BASED ANALYSIS OF QUERY LANGUAGE FOR DIRECTORIES AND APPLICATION

Mr. Thieu Quang Trung

Prof. Phan Minh Dung

LDAP (Lightweight Directory Access Protocol) is a protocol for online directory services accessing. It provides a standard model and protocol used in today's modern directories. Recently, the XML (eXtensible Markup Language) is rapidly becoming the globally accepted specification for exchanging complex structured data. In LDAP, data are organized as a tree, where each node can contain data value and can act as a namespace for other nodes. This corresponds closely to XML, since the XML data model is hierarchical in structure and usually implemented by considering the XML document as a tree structure. This thesis studies the expressive power of LDAP query language and the use of LDAP to store XML data. The study of the thesis also considers the applying of storage Windows Registry data in LDAP data model.

CS-03-08

DEVELOPMENT OF CONTEXT-BASED KEYWORD EXTRACTION TOOL

Ms. Rewadee Sukatipan

Prof. Ramakoti Sadananda

Current approaches to information retrieval over the Web are based on a scenario in which the user enters a query to a search engine. The search engine then retrieves an ordered set of documents that best match the user's query. This thesis proposes the basic setting of the search scenario by introducing context as an additional input of the search process. In this scenario, when the user marks an expression in a document and submits it for search, the system captures the context surrounding the expression, and utilizes it to yield focused results. Using the context to guide the search constitutes a considerable algorithmic challenge. One needs to find ways to extract the amount of context that optimizes the information retrieved, as well as to devise adequate ways to use the extracted context for focusing the response to the user's query.

CS-03-09

EVALUATION SYSTEM FOR EMBEDDED CONTROLLERS

Mr. Truong Quang Vinh

Dr. Nitin Afzulpurkar

Nowadays, embedded systems with Linux OS are commonly applied in automotive industry. Some of applications require strict time response, and others need to be exactly scheduled to execute a period task. All of these are called time-sensitive applications. Measuring and evaluating time parameters of an embedded system for time-sensitive applications is very necessary for developers to guarantee that it works functionally. For this purpose, signal stimulating and signal analyzing hardware are required. In this thesis, modular signal generating and signal analyzing means will be designed. The prototype of the evaluation system will be implemented to test and verify an embedded system. Improving the performance of the system is

the next part of this thesis. The research focuses deeper on the possibility of turning Linux into a real-time operating system. Particularly, it investigates available real-time solutions for Linux, but also looks into the soft variants for completeness. A number of performance tests are conducted during the improvement to make sure that the implementation meets the demands of a real-time operating system.

CS-03-10

AN ADAPTIVE ASSESSMENT SYSTEM FOR E-LEARNING

Ms. Nguyen Thi Bich Nga

Dr. Peter Haddawy

Appropriate feedback and improvement support are essential factors that constitute the effectiveness of an assessment system. This study proposes an adaptive assessment system that aims to enhance those factors. For each module of knowledge tested, a meaningful feedback is reported which points out exactly which fundamental components included in the module is currently at risk. The system then recommends student review materials and offers multiple traits adaptive assessment, that has ability to focus on improving weak points, as an efficient way to help student improve their performance in these components. Current assessment systems do not support tracking and analysing information as well as provide feedback and tailored support for fundamental components included in assessed module. Due to this fact, the system is proposed as a way to make performance improvement process of student in E-Learning more efficient.

The major challenges of this study have been to point out exactly the risk fundamental components and generate a multiple traits adaptive assessment that content focus on improving these weaknesses. The contribution of this study is an E-Business course with embedded adaptive assessment system that can be used over Intranet or Internet. ASP.Net technology and SQL Server are used in implementation.

CS-03-11

COMPUTER VIRUS CLASSIFICATION IN the WINDOWS OPERATING SYSTEM

Mr. Vo Mau Pha

Prof. Phan Minh Dung

Computer Viruses have been existing twenty years long. There are thousands of them in the wild. Many of them have common characteristics as a real creature: life time, environment and their evolutionary. Their successful spreading depends on not only in the computing environment but also in the mechanism built in it.

One important task that helps human being understand and live with the world is to classify thing into sub domain. In this thesis work, we do the same thing about computer viruses: Classify them based on their characteristics, environments and behaviors. Once we have a clear classification about them in one environment, the similar task is nearly finished to the other because the process is the same, we know where to look for the computer virus problem.

CS-03-12

DETECTING POSSIBLE BUFFER OVERFLOW ATTACKS IN C PROGRAM

Mr. Vuong Hoai Nam

Prof. Phan Minh Dung

Buffer overflows are unintentional accesses to memory outside the intended object. In recent years, buffer overflow attacks have become the most prevalent techniques used to hack both remote and local computer systems. All the reason for such a profound spread of this kind of attack is that many software and operating systems are written in C language. C is inherently unsafe because array and pointer references are not automatically bounds-checked. Therefore, it is up to the programmers to do it themselves. Nonetheless, these security holes are often wrongly omitted or checked. In addition to its prevalence, this attack can cause the most serious problem to compromised systems. One solution to buffer overflow attacks is to statically detect possible vulnerabilities in C source codes before they are deployed. Much research has been done to solve this problem, however, in a subset of C features, which involves pointer, array of char, one-dimension array, pointer arithmetic statements, and standard C string functions.

This thesis will examine the principles of Unix/Linux buffer overflow vulnerabilities. Then, it will survey the existing static methods of detecting buffer overflows, and propose a better method that accumulates a larger domain of problems involving arrays of pointers. Finally, a prototype is implemented based on the proposed method.

CS-03-13

FORMAL VERIFICATION OF FAIRNESS PROPERTIES IN SECURE ELECTRONIC TRANSACTION (SET)

Mr. Nguyen Trung Hieu

Prof. Phan Minh Dung

Secure electronic transaction protocol is an important security protocol used to secure bankcard payments on internet. Beside authentication, authority & privacy requirements, this protocol needs to meet fairness requirements that ensure no protocol participant gains advantage over another. Non-repudiation is one of the most important properties in fairness requirement. Non-repudiation guarantees that merchant can not deny having sent the product and customer can not deny having received it. Existing version of SET does not support non-repudiation property when good is digital like software's.

This thesis proposes a new protocol that is based on original SET protocol and satisfies this requirement. To prove this new protocol satisfies fairness properties, Game based method proposed by Kremer is applied. This method uses alternating

transition system to model protocol and alternating time temporal logic to express requirements. This method is automated by using the model checker MOCHA. However, in my verification process I find that using MOCHA tool directly is not correct. MOCHA tool verifies formula at initial state only, and lacks ability to verify formula at certain reachable state of system. In this thesis I have put efforts in direction to enhance the capability of this tool also.

CS-03-14

DEVELOPMENT OF A WEB-BASED NETWORK MONITORING TOOL

Mr. Nguyen Nhat Tan

Dr. Peter Haddawy

Raw packets that pass the Proxy Sever or Bastion Host contain information about the incoming and outgoing traffic of network users. The information includes kinds of traffic (e.g. HTTP, FTP, SMTP), source and destination IP address, traffic level by traffic type and users (based on IP addresses). The raw packets are collected and analyzed by network monitoring tools to support network management.

Currently, there are several such software tools available, but they usually collect real time network data only, so network managers cannot get the network data from the past in order to analyze long-term network usage. Also many of these tools do not provide a user-friendly web interface. Thus, they cannot be used to monitor the network from the Internet or intranets. This makes them very inconvenient for network managers.

This thesis develops an engine that captures raw packets from a network and stores them into a database for long-term network traffic analysis. A web interface shows network usage in a user-friendly presentation to network managers. This tool is limited to monitoring a local subnet with unicast IP traffic and Ethernet topology.

CS-03-15

A CERTIFICATE-ISSUING SYSTEM FOR THE UNIVERSITY OF AGRICULTURE AND FORESTRY OF VIETNAM

Ms. Mai Anh Tho

Prof. Phan Minh Dung

With billions of users on the Internet, more interactions will occur between strangers, i.e., entities that have no pre-existing relationship and may not share a common security domain. Before strangers can conduct sensitive interactions, a sufficient level of mutual trust must be established. Traditional security approaches based on identity within the local security domain now become less efficient and reliable. Users are unknown at the time entering system so that the system can not store identities before and then it does not determine which set of roles is mapped to users.

Automated trust establishment offers a solution to this practical problem. With automated trust establishment, strangers establish trust by exchanging digital certificates, the online analogues of paper certificates: digitally signed assertions by a certificate issuer about the certificate owner. Public-key certificate is one of the most important types of digital certificate, in which a public-key value is associated with a particular person or entity. A public-key certificate is signed using the issuer's private key, and can be verified using the issuer's public key. Now, the authorization, which consists of authentication based on trusted public-key certificates and direct assignment of a set of corresponding roles, is more secure.

In this thesis, we study encryption, message digest, signature algorithm and public key certificates in terms of implementing a prototype of a system that issues public-key certificates. University of Agriculture and Forestry of Vietnam is chosen as a case study of such system and public-key certificates have instances as electronic degrees for graduated student of the university.

CS-03-16

CIRCULAR BIT STRINGS AND STABLE PERMUTATIONS

Mr. Nguyen Ngoc Trung

Dr. Sumanta Guha

In this thesis we consider the problem of recognizing and generating permutations of the integers $\{0, \dots, 2^r-1\}$ that have a special property. The idea is as follows. Any integer from $\{0, \dots, 2^r-1\}$ can be represented in its binary form as a bit string of length r . Consequently, a permutation $p = (p^0, p^1, \dots, p^{2^r-1})$ of $\{0, \dots, 2^r-1\}$ can be represented as a bit string, call it $R(p)$, of length $r \cdot 2^r$ by concatenating the binary representations of integers in the permutation. Imagine now the bit string $n(p)$ to be arranged circularly. Our question then is, to obtain a permutation of $\{0, \dots, 2^r-1\}$, can we start reading of binary representations of length r , from any bit on this circle? Clearly, by our definition of $R(p)$ there is at least one such valid start bit. The permutation p is said to be stable if any bit of $R(p)$ (arranged circularly) could serve as the start bit for finding a permutation of $\{0, \dots, 2^r-1\}$.

In this thesis we prove several interesting properties of stable permutation. Algorithm for recognizing stable permutations, rules to generate stable permutations are also proposed.

CS-03-17

DETECTING POSSIBLE FORMAT STRING OF AGRICULTURE AND FORESTRY OF VIETNAM VULNERABILITIES IN C PROGRAMS

Mr. Huynh Buu Ky

Prof. Phan Minh Dung

In June 2000, the first public release of format string vulnerability against wu-ftp 2.6.0 has obtained major attention. This vulnerability arises from the combination of unchecked variable argument (varargs) functions and standard C library

implementations. Since then, various format string exploit techniques have been discovered and numerous incidents of format string attacks have been reported to date [1, 2, 6]. The exploitation of format string bug represents a whole new serious class of vulnerabilities in C programs that can be used to gain highest privileges on a local or remote host.

The first part of this study explains the nature and analyses the tricks and limitations of format string exploits. A semantic model of simplified version of printf function is also established to give a formal explanation how format string bugs occur and how they cause the change of program flows. Part two discusses currently well-known static and run-time defense strategies against this important class of vulnerabilities. Finally, we present our system, which combines both static analysis and run-time checks. The system employs bottom-up qualified type inference engine to perform static analyses to check source code for safety printf-like function calls, and automatically inserts run-time checks where safety cannot be guaranteed statically. Our tests on several vulnerable source code show that our system exhibits a lower rate of false positive than current top-down qualified type inference one. Moreover, performance test shows that the system imposes lower performance overhead than currently run-time solution.

CS-03-18

COMPUTER VIRUS CLASSIFICATION IN UNIX ENVIRONMENT

Mr. Le Tan Phuoc

Prof. Phan Minh Dung

In just over a decade, most of us have been familiar with the term computer virus. Even those of us who don't know how to use a computer have heard about viruses through Hollywood films such as Independence Day or Hackers and some means like that. International magazines and newspapers regularly have virus-scares as leading stories. There is no doubt that our culture is fascinated by the potential danger of these viruses.

Computer virus have become threat to computer users and almost every field in the advance technology industrial nowadays. Know about virus and environments that can be infected by virus is very necessary for anti-virus researchers as well as operating systems makers. With the development of the open source systems today, computer viruses on these systems should be considered strictly. This study is about Unix environment, from the analysis of internal mechanisms of Unix, propose some viruses that can work on this environment and suggest methods to prevent as well as restrain damages of these viruses.

CS-03-19

LOGIC PROGRAMMING BASED TRUST MANAGEMENT

Mr. Tran Ngoc Thai Kha

Prof. Phan Minh Dung

Trust management (TM) problem is a promising approach for authorization and access control in distributed systems according to given security policies. Aspects of TM include formulating application-independent security polices and security credential, determining whether particular set of credentials satisfies the relevant policy, and deferring trust to third parties. The core of TM consists of policy language and policy engine.

In this thesis, we clarify a well-known Trust Establishment system (Herzberg et al., 2000) by transforming TPL+ language and engine to logic programming context. It is concluded that its solution for negation is very simple and so weak. After that, we design and implement a Trust Management system based on logic programming. The policy language of this system is normal restricted logic language and the policy engine is extended from Fitting semantic and Stable model semantic. This algorithm is proved to be sound for Preferred extension semantic.

CS-03-20

APPLYING GRID TECHNOLOGY TO E-COMMERCE: A CASE STUDY IN SELLING SOFTWARE AND APPLICATIONS

Ms. Phan Phuong Lan

Prof. Phan Minh Dung

Grid technology has gained attention over years. It contributes significantly in sharing flexibly and securely resources among dynamic collections of individuals and institutions. This technology is being applied widely in science communities where an institution can freely share and access resources with/to other institutions, but it has not yet been applied in the commercial world where the needs of sharing resources are large. In that world, customers want to save cost and time from buying resources and/or resource accesses on demand; providers want to make benefits from selling the resources and/or resource accesses. All current systems that support buying and selling resources has just focused on buying and selling software. Many systems support buying and selling resources accesses, but these resources are web applications. No system allows customers and providers to buy and sell resources accesses that are hardware. So that the objective of this thesis is to build a system that supports buying and selling online resources demanded by customers. The resources are software, hardware accesses and applications. The Globus Toolkit, the well-known toolkit of Grid technology, is used in building the system. The designed and implemented system not only supports possibly customers to save their time and money by using brought software and/or rented hardware accesses and applications, but also facilitate providers to obtain benefits from selling software and/or sharing their hardware and applications.

CS-03-21

MULTI-AGENT VIEW OF DETECTOR AND CORRECTOR: A CASE STUDY IN DISTRIBUTED MINIMUM SPANNING TREE

Ms. Narisara Boonruang

Prof. Phan Minh Dung

The study of self-stabilization has been very attractive to computer scientists over the past 25 years. Intuitively, self-stabilization is the ability of the system to be initialized in any state of the system, and converge to its desired behavior within finite amount of time.

The concept is very important to large-scale heterogeneous systems with autonomous components, which can be abstractly called multi-agent systems. These systems exist everywhere; for example, the Internet, cellular and PDA communications, international trade, multi-national corporate databases, multi-user games, and even local area network.

However, there is no effective general model to describe self-stabilizing algorithms. Such a model will make it easier to understand self-stabilizing algorithms, and to design a new algorithm.

This thesis selects the theory of detector and corrector, which describes fault-tolerant systems as components, to be applied into self-stabilization concept. The multi-agent system is focused. Therefore, the distributed minimum spanning tree algorithm is selected as a case study.

The thesis aims to study how well the theory can describe self-stabilization through the view of the selected case study. Also, it aims to study how each agent in the multi-agent system participates to create components as described in the theory.

CS-03-22

AN ONTOLOGY MAPPING FRAMEWORK

Ms. Jittiya Kaewprag

Prof. Vilas Wuwongse

The Semantic Web becomes ever more important issue for the Web today. It relies strongly on formal ontologies to structure data for comprehensive and exchangeable machine understanding. Furthermore, information on the Web is naturally distributed. Each information source uses its own ontology to represent its data. An interoperability problem among different information sources is therefore a challenge issue on the Semantic Web. One way to overcome that problem is to provide a systematic way for mapping and merging ontologies. However, there are some limitations for recent related research. Most approaches are usually based on heuristic or syntactic clues to determine equivalence between ontologies. Thus, a new ontology mapping framework is proposed by applying a quantitative technique called semantic similarity measure to compute a similarity value between two concepts. The semantic similarity measure can be retrieved from properly combining three semantic factors consisting of the shortest path length between two concepts, the depth and the information content of a subsumer by using WordNet, an online lexical dictionary, as a semantic reference source. Besides, a new concept is also introduced, called the weight-based approach, to support strong semantic relation word pairs by adding weight into semantic similarity measures. Consequently, this technique displays a dramatic improvement over the previous research. Based on the proposed approach, the automatic financial report mapping system is developed as a demonstration system of its practical usage in a real-word domain.

CS-03-23

SIMULATION OF TRADE IN BARTER TRADE EXCHANGES

Ms. Khaimook Dhananaiyapergse

Dr. Peter Haddawy

Modern barter business, a kind of B2B e-Marketplace, is a trade exchange done via Web-based technologies between two or more companies associate with a broker company. This kind of trade grows rapidly in the past few years.

There are researches on B2B e-Marketplace concentrated on improving the system efficiency, such as customer purchase prediction, system trade volume maximization and system trade balance optimization. However, those researches have never been observed their properties and effectiveness against the real system. Thus, this thesis contributes to the issue.

Main purposes of this thesis are to learn the purchase behavior of companies in a barter pool and simulate the purchase behavior of companies. We apply Naïve Bayes classifier with aspects as purchase the behavior learning. The inverse-transform method is conducted to simulate the learned purchase behavior.

Experiments conducted are divided into three parts. The first part is designed to evaluate the learned purchase behavior. The second part is designed to evaluate the simulated purchase behavior. The last part is designed to evaluate the optimization algorithm. The evaluation results of the first two parts show that the purchase behavior learning algorithm and purchase simulating algorithm are able to accomplish their objectives. The evaluation result of the third task shows that the optimization module is able to reduce the system absolute balance and increase the system trade volume.

CS-03-24**NON-REPUDIATION AND FAIR ELECTRONIC PAYMENT SYSTEMS**

Ms. Chonlatorn Chunhavikosit

Prof. Phan Minh Dung

Without non-repudiation and fairness property, the electronic payment (e-payment) systems do not have a chance of acceptance from users. This thesis proposes an e-payment protocol regarding non-repudiation and fairness properties called 'fair 3KP'. The protocol is modified from iKP protocol, which is one of a secured e-payment method for credit card-based system. The prototype system of fair 3KP is implemented to test its applicability in real world situation. A trusted third party (TTP) is the important component to make fair 3KP to be practical and therefore its trustworthiness is analyzed. The guideline for establishing trustworthy TTP service in Thailand is proposed to enhance the feasibility of fair 3KP in a real use.

CS-03-25**A WEB SERVICE BROKER FOR NON-EXPERT CLIENTS**

Ms. Orawan Meeporncharoenkit

Prof. Vilas Wuwongse

The Web services paradigm enables automated web application interoperations. The first step toward this interoperation is the location of other services. Although there are many service brokers performing location of Web services, most of them do not allow a client to search for Web services by their capabilities. The traditional UDDI search is restricted to keywords and key reference matching. Moreover, with the evolution of Web service technology networked services will not only become increasingly sophisticated, but also move into the area of business-to-consumer interactions. This study suggests a Web service broker locating a service based on not only service capabilities but also user preferences. The broker employs a unified modeling language XML Declarative Description (XDD) and an XML-based declarative programming language XML Equivalent Transformation (XET) to perform a user's preference based matchmaking, which is easy and useful for a non domain-knowledge user. The prototype system is constructed as a domain-independent framework, therefore a computer business is one possible application applied on the broker.

CS-03-26**UML MODEL CHANGE DETECTION AND MANAGEMENT FOR CONSISTENCY MAINTENANCE**

Mr. Nimit Pattanasri

Prof. Vilas Wuwongse

In UML-based software development, user requirements are often changed in any phase of development life cycle. This causes not only the program code but also the design, in particular UML diagrams, to be changed. Consequently, the change may introduce an inconsistency between UML diagrams. The thesis proposed a practical system to detect the inconsistency in an incremental way. In addition, the system provides more precise repair actions to a user.

Three main parts of the system are described, that is, change detection, change inference, and model merging. Lightweight logical change detection is proposed to find the changes in two versions of a UML model from a user. Its main task is to translate physical changes into logical ones. Later on, the inference engine, XET, with consistency ECA rules checks whether the user changes cause inconsistency. If so, the system provides appropriate actions (or recommendations) to resolve the problem. Finally, the system applies all changes including inferred ones, needed to maintain the consistency, to the original model of a user.

CS-03-27**SLIDING WINDOW ASSOCIATION RULE MINING**

Ms. Onanong Nopkhun

Dr. Peter Haddawy

In rapidly thriving barter exchange business, human brokers as business mediators take care of all their members in various ways such as what they will purchase, what products or services will be provided, and purchase recommendation. For recommendation, brokers need to observe all their members' purchase behavior. How effective the recommendation is depends on an individual broker's experience. In order to help new brokers generate useful recommendation in a short time, a recommendation engine, a semi-automated broker, is proposed in this thesis. The objective is to observe and predict purchases based on sparse real business transactions.

Among several data mining techniques, a novel approach, Sliding Window Association Rule Mining (SWARM) algorithm, is proposed to discover the association rules from temporal transactional data. The correlations of products purchased within a given time window are considered as the rules that are used for recommendations at appropriate thresholds. In the experiment, three parameters: minconf, minsup and winsize, were used to find out the affect of these parameters on the association rules.

The experimental results of testing with real data show reasonable and accurate association rules. The results were evaluated via an evaluation program and a human expert. First, the evaluation program shows that small time window is found to be slightly more helpful than large time window. Second, the human expert who classified how useful or interesting these rules are shows that there are still some rules relatively useless to generate recommendations. As one of the confidence measures, lift is appropriated for ranking top-N rules and also improving the generated rules to become more efficient for recommendations.

CS-03-28**WEB SERVICES DEVELOPMENT BASED ON ANALYSIS FOR CONSISTENCY MAINTENANCE AND DESIGN PATTERNS**

Ms. Pabita Dhakal
Dr. Dencho N. Batanov

Web services technology are the latest technology in the application development. They provide new business opportunities due to many advantages such as interoperability, just-in time integration and wide industry support. Software patterns are mechanism to document and share design solution in developing software. Since the Web services have its own features and standards, it is not useful to rely on the existing software patterns and development methodology. In this study Web services patterns are identified and illustrated as well as methodology to develop web services based on patterns is described. The proposed methodology covers overall lifecycle of development. This methodology covers analysis, design, implementation, testing and publishing phases. It is based on Web services patterns along with modeling techniques and framework to design service implementation. This helps in effective creation of Web services.

The proposed methodology is used to develop Web services in the QandA teaching/Learning system. It is tested by developing client application for correctness and validity. This study also shows transition from traditional Web application to Web services based application.

CS-03-29

A UNIFIED FRAMEWORK FOR VERSION AND ACCESS CONTROL OF XML DOCUMENTS

Ms. Kantinee Katchapakirin
Prof. Vilas Wuwongse

As more and more documents are made available in the eXtensible Markup Language (XML) format, concerns are being raised about XML document management problems. Existing research has investigated access control and version control functions and their various extended models separately. Because of a strong correlation between both functions, the most important issue, the efficient management of interrelation, remains open.

This thesis introduces a unified framework integrating version and access control techniques in an efficient way to realize advanced management of XML documents. The main contribution of this research is to present a method to automatically assign predefined instance level policy documents with a new version of protected documents by analyzing changed content. Moreover, by extending the framework with temporal features, the utilization of the temporal data can be supported. For example, the framework has the ability to retrieve the documents, which are valid at a certain time.

Furthermore, the unified framework is also designed to realize document management system interchangeability, interoperability and integrability by subdividing a foundation layer into unified representation and computation sub-layers. The framework is employed by a unified modeling language, namely XML Declarative Description (XDD). XDD not only directly employs HL7 CDA and XACML standard as a protected document and access control policy representation, but also enables direct computation of all functions and flexibility of knowledge modification, which characterizes the framework. The separation of foundation introduces more flexible system architecture, so that different vendors can use their own computational models.

CS-03-30

FLEXIBLE WEB SERVICES COMPOSITION

Ms. Bossaporn Eampornchai
Prof. Vilas Wuwongse

A web service composition is categorized into two types, that is, static and dynamic web service composition. The static one selects the services to be composed at design time. The dynamic one selects those at run-time. The advantage of dynamic composition is well recognized. For example, it yields high quality services composed together. However, after finished choosing these services, the composition is just the same as a static composition.

That is, the structure of the composition is fixed and not easy to change. This thesis proposed a framework for a flexible web service composition, whose structure is very general so that it satisfies a variety of user requirements.

CS-03-31

OPTIMIZING IN B2B E-MARKETPLACE ECONOMIES

Ms. Namthip Rujikiatkamjorn
Dr. Peter Haddawy

In this thesis, we investigate various techniques used by trade brokers in barter trade exchanges in order to maximize trade volume and ensure equitable participation of members in trade. In particular, we propose mathematical models to solve the problem of maintaining balance of trade and fair distribution of trade, which maximizing single-period trade volume. We show how the trade balance and fairness problems can be represented in terms of network flow models and propose an efficient algorithm for the case of uniform product cost or when fractional solution are acceptable. For the case where integer solutions are required for non-uniform product cost, we present two rounding algorithms that take the fractional output of the network model and produce an integer solution. We use a linear, non-linear, and integer programming solver (Lingo software) in order to verify the solution of our algorithm. We empirically compare solutions from the network flow with rounding approach and mixed integer programming approach. We evaluate the effectiveness of our optimization procedure in maintaining trade balance by using a simulator built using transaction history data from a trade exchange. We also measure the effect of trade balance optimization on the long-run trade volume. Our results confirm the barter trade exchange rule of thumb that maximizing single-period trade volume while maintaining balance of trade helps to maximize trade volume over the long run.

CS-03-32

JOINT SEPARATION OF CLUSTERS: THEORY AND USING PHP EXPERIMENTS

Mr. Pattarawit Polpinit

Dr. Sumanta Guha

In this study we investigate a measure of the so-called *joint separation* of clusters from both the experimental and the theoretical points of view. In the experimental part, we measure the *joint separation* of a finite set of planes F of cardinality $n \geq 2$ in term of the *minmax* angle of vectors on the planes belonging to a given set. *Minmax* angle is defined as the largest angle θ such that, if n vectors are chosen, one each on the plane, then at least two of the vectors have an angle of at least θ between them. Since obtaining *minmax* angle for arbitrary finite sets of plane in R^3 seems hard, we restrict to planes bounding the faces of a regular polyhedron. An algorithm based on Hill Climbing and Lingo model is used in the search for *minmax* angle of regular polyhedron. In the theoretical part, we study the generation of *minmax* angle as a notion of *joint separation*. Given a set of points of various colors on the line, an interval is called *color-spanning* if it contains at least one point of each color. We present an efficient algorithm to solve a problem of finding the minimum *color spanning* interval. Furthermore a semi-dynamic and a fully dynamic algorithm to maintain such interval are proposed.

CS-03-33

COMPARATIVE ANALYSIS OF C++, JAVA, AND C# RELATED TO WEB SERVICES DEVELOPMENT

Mr. Pichet Premteerawatchai

Dr. Dencho N. Batanov

In the world of web technology, web services are a new technology. Web Services have many components; for example, SOAP (Simple Object Access Protocol), WSDL (Web Services Description Language), UDDI (Universal Description Discovery and Integration), and XML (eXtensible Markup Language). Web services provide services that have specific functions for web clients. Developers can choose many object-oriented programming languages for building web services. Popular languages in web services are C++, Java, and C#. In this study, a comparative analysis of C++, Java, and C# related to web services development will be conducted. Also, calculator a web application, which is a simple web service, will be created. The environments of three languages, which are use for creating web services, is gSOAP compiler (for C++), J2SE and JWSDP (for Java), and Visual Studio.NET (for C#). The entire processes of developing calculator web services are described throughout this study.

INFORMATION MANAGEMENT

IM-03-01

A CASE STUDY OF BAYESIAN NETWORK THEORY REFINEMENT FOR PREFERENCE ELICITATION

Ms. Rachanee Srisurangkul

Dr. Peter Haddawy

This thesis explores the use of one particular theory refinement technique, Bayesian networks (BN), to learn user preferences. Bayesian networks are in efficient and effective representation of probability distributions via conditional independence. We demonstrate this approach through the example, which involves preference under certainty. The initial network is derived directly from a domain theory of propositional Horn-clause rules that encode assumptions concerning preferential independence. The network structure and parameters are then refined by training on data representing an individual's preferences. We empirically compare the Bayesian network with a neural network approach in terms of learning rate and accuracy.

IM-03-02

DEVELOPMENT OF USER-CENTERED WEB SITES WITH HIGH USABILITY

Mr. Shahab Alizadeh

Dr. Dencho N. Batanov

Usability concept has become the hot topic and the critical aspects of many online applications. Many web-based application with great functionality has failed just because of ignoring the term of usability. Usability has many aspects but for web based application mainly is focused on GUI and functionalities. This study a methodology of usability pattern to give the designer the first needed steps in reaching to better GUI has been presented.

Numbers of usability functions based on the framework technology are build in order to be integrated with the Question and Answer model by the same time keeping the concept of object-oriented terminology of the system. Object-oriented software development life cycle is been applied for analysis and design of the application. Finally implemented functions are applied in the Q&A system for representing their functionalities and tasks.

IM-03-03

MAPPING WORLD POVERTY USING A SELF-ORGANIZING MAP

Mr. Nguyen Van Thuan

Prof. Ramakoti Sadananda

The Self-Organizing Map (SOM) is a popular neural network model. The SOM quantizes the data space formed by the training data and simultaneously performs a topology preservation of the data space on a regular two-dimensional grid. The SOM also has an excellent visualization capability including techniques to give an informative picture of the data space, and techniques to compare data vectors or whole data sets with each other. The SOM can be used for clustering, classification and modeling. In this research, the SOM technique is applied on mapping the poverty data available from countries. The world poverty map is based on multi-dimensions of poverty taking into account a number of indicators. The map groups countries of similar levels of

poverty together, thus providing a visualization of structure of poverty. The data used for this study was extracted from the World Development Indicators (WDI) published by the World Bank. Ten poverty indicators were selected from the WDI, which related to quality of life, health, education, and sanitation. The map in this study is obtained through self-organization cluster poverty in 132 countries into 4 groups. Two-dimensional representations of the data are presented showing the countries in each group.

IM-03-04

A LOAD SHARING STRATEGY FOR DISTRIBUTED WEB SERVER SYSTEM

Mr. Truong Quang Khai

Prof. Ramakoti Sadananda

A distributed multi server Web site can provide the scalability necessary to keep up with growing client demand at popular sites. Load balancing of these distributed Web server systems, consisting of multiple Web servers for document retrieval and a Domain name server (DNS) for address resolution, opens interesting problems. Unlike traditional parallel/distributed systems in which a centralized scheduler has full control of the system, the DNS controls only a very small fraction of the requests reaching the multi server Web site. This peculiarity, especially in the presence of highly skewed load, makes it very difficult to achieve acceptable load balancing and avoid overloading some Web server. DNS plays a very important role for balancing the load on multiple servers of cluster.

Load balancing on multiple servers ensures that Web traffic doesn't overload one server while other servers sit idle, especially, in an environment in which the access frequencies from different client domain are not evenly distributed and the Web servers have different capacities. The point is how to choose the best candidate server location for retrieving data. The main focus of this paper is to study of DNS from this point of view.

IM-03-05

RELATIONSHIP BETWEEN HIERARCICAL ORDER PLANNING AND BOTTOM UP PLANNING

Mr. Nguyen Anh Hao

Prof. Phan Minh Dung

Linear time temporal logic has been used as a tool to express control knowledge in reactive systems. Its attractive potential is proved by TLPlan, which is being considered as the fastest planning system. However, the design of TLPlan has a drawback due to the use of pruning method which limits the efficiency of the system thus whether the control knowledge in TL control formulas can be used for selecting plan operator has been pursued in recent years.

SHOP has similar potential with TLPlan in expressing control knowledge so as it can express TLPlan's control knowledge. The primary reasons for converting control knowledge of TLPlan into SHOP are (1) TLPlan system is based on state space search while SHOP uses problem deduction and it has long been known that problem deduction is much more efficient than state space search, (2) The use of temporal logic control formula in TLPlan has been focused to state pruning rather than actual planning.

In this thesis, the transformation is described for a sub-class L1 of TL formulas, in which state constraints generated by a TLPlan's control formula in each pruning step are all re-established in each task selection step of equivalent SHOP planner.

IM-03-06

OPTIMIZATION ALGORITHMS IN BARTER TRADE EXCHANGES

Mr. Nguyen Phat Tai

Dr. Peter Haddawy

Barter is a fast-growing, multi-billion dollar industry. The core business of a barter trade exchange is to match buyers and sellers; this task is currently done by human brokers. Brokers attempt to manage the trade within a barter pool of companies in order to maximize various economic objectives. This thesis aims at automating the economic optimization aspect of the brokerage process. A barter pool is modeled as a matrix, where rows represent member companies and columns represent products and services. Entries indicate whether a company sells or wishes to buy a given product or service. Optimizing trade within such a matrix amounts to a combinatorial optimization problem. Two versions of the problem are investigated: a simplified qualitative version and a quantitative version. For both types of problems several different optimization algorithms are developed and evaluated, including exhaustive search, heuristic search, and nonlinear programming. Experimental results on synthetic data show that heuristic search and nonlinear programming yield optimal solutions for the majority of problems and that when they return sub-optimal solutions, the solutions are typically very close to optimal.

IM-03-07

EFFICIENT HEURISTICS FOR OPTIMALLY MATCHING BUYERS AND SELLERS IN E-MARKET PLACES

Mr. Bui Cong Giao

Dr. Peter Haddawy

This thesis addresses the problem of matching buyers and sellers in barter trade exchange e-marketplaces. A barter trade exchange is a collection of businesses that buy and sell products among themselves. The collection of businesses is viewed as a micro-economy, so that matching is viewed from an economic perspective. An optimal matching seeks to maximize trade volume and to ensure that all companies share in the trade. The matching problem is given a formal representation and an efficient heuristic search algorithm is developed to solve it. The quality of solution of the heuristic search algorithm is evaluated by comparing it to the optimal solution obtained by exhaustive search on a large set of problems. The algorithm is shown to be fast enough to deal

with very large real-world problems. The developed technique has the potential to greatly benefit the barter trade exchange industry as the size of trade exchanges grows.

IM-03-08

IDENTIFICATION OF WEB COMMUNITIES

Mrs. Nguyen Thi Minh Huong

Prof. Ramakoti Sadananda

This research is about the identification of a web community. A community is known by the higher interaction among its member, than between members and non-members. The interaction is expressed in the form of links. The identification of a web community would help to increase the efficiency of search and improve the precision of the result. The focus here is to use the maximum flow network algorithm for the purpose. The thesis also reports the development of tools for the identification of the web community and an implementation to test the tools.

IM-03-09

CRM DECISION SUPPORT SYSTEM FOR MULTI-SERVICE/PRODUCT PROVIDERS

Mr. Thitipong Navalertporn

Dr. Vatcharaporn Esichaikul

In recent years, businesses are increasingly realizing the significance of knowing their customer better. More and more companies are focusing their efforts on building long-lasting ties with customers; therefore, Customer Relationship Management (CRM) is becoming a priority for many businesses.

This study concerns with designing and implementing a prototype Decision Support System for Customer Relationship Management in order to enhance an effectiveness of marketing operation in the Steel Manufacturer and Distributor Company. In recent years, steel business has become one of the most fiercely competitive industries. In such competitive environment, traditional customer relationship strategies of marketing department deployed to fuel business growth become less effective. This study, therefore, proposes two primary applications, which include (i) customer segmentation and (ii) customer retention analysis. The objective of customer segmentation is to define the right strategy for the right customer. Customer retention analysis, on the other hand, is to identify the customer who is likely to leave or drop the services. Once the prototype system is developed, an evaluation on the accuracy of the results is conducted. Finally, some possible areas for the future development of a decision support system for customer relationship management are suggested.

IM-03-10

PREDICTING PRODUCT PURCHASES FROM TRANSACTION DATA USING ASPECT MODELS

Mr. Ha Hong Thuy

Dr. Peter Haddawy

Microsoft .NET is the latest generation of software technologies that makes it easier to share information between computer systems. It allows web development giving the developer a notion similar that of developing a desktop application. It is a generation ahead of its predecessor ASP in terms of architecture and performance. ASP.NET has tried to work on many of the shortcomings of ASP and added features like object-orientation and multiple language support; it has also evolved into a compiled module. To take advantage of this new technology one can transition existing ASP applications through a prescribed strategy and a set of guidelines. The resulting guidelines aim to provide a feasible transition to the new environment with minimum amount of workload, such that maximum amount of resources could be utilized from the present system to work on the new environment. The prescribed strategy and a set of guidelines have been implemented on the Q&A e-learning system.

IM-03-11

A MODULAR APPROACH TO E-LEARNING CONTENT CREATION AND MAINTENANCE

Mr. Phan Thanh Duc

Dr. Peter Haddawy

There are thousands institutes, universities using e-Learning as a formal method in training and there are millions people using e-Learning as a principal method to study. In addition, there are thousands institutes, universities, companies working as course content developers to provide e-Learning resources to the learners.

However, the content of e-Learning courses is not static. They are changed with the evolution of the society. Therefore, the management of the course content are challenging and demanding tasks. One of them is how can update the course content in easy way without breaking through the existing course structure. Nevertheless up to now, such techniques and tools for updating course content are not available in the market.

This thesis will involve in developing technique for creating modular multi-media e-learning materials that can be easily updated based on Sharable Content Object Reference Model (SCORM) – a standard of e-Learning. The effectiveness of the technique will be tested by preparing material for a course on E-commerce that will be taught to students at AITCV (Asian Institute of Technology – Center in Vietnam).

IM-03-12

MULTI-LANGUAGE E-DICTIONARY SUPPORT FOR WEB-BASED TRAINING

Mr. Do Vinh Truc

Dr. Peter Haddawy

E-learning has been becoming more important in information technology. Not only utilization of courses in E-learning, but also helping students for studying is one of requirements today. This thesis builds a program that is one of three parts in our E-learning project that can satisfy the question. We build the program as a web tool that allows instructor applies for creation of definition, and insertion a hyperlink to a keyword that user choose on web browser. Based on data in e-learning course, we build a tool which an interface like dictionary on the web. Hundred of definition in Electronic Commerce is included in the program, and an English-Vietnamese dictionary about 53000 words is also embedded.

IM-03-13

PERSONALIZED LOCATION BASED TOURISM SERVICE FORMOBILE USERS

Ms. Wang Hua

Dr. Vatcharaporn Esichaikul

Mobile users such as tourists require mobile services. Handheld devices promise access to a range of travel-related services while on tour. A number of usability issues, however, still ask for intelligent new solutions. Issues are, for example, the limited capacity of handheld devices, user interaction via small screens, mediation of heterogeneous services, as well as other issues of wireless access to internet-based services. They offer new opportunities for mobile commerce application and services to end-users, mobile network operators continuously seek new and innovative ways to create differentiation and increase profits. One of the best ways to do accomplish this is through the delivery of highly personalized services. One of the most powerful ways to personalize mobile services is based on location.

The thesis constructs a personalized location based services model for tourism in order to improve the tourism services. In the implementation part, a prototype of restaurant recommendation system was developed. The proposed system provides a restaurant recommendation by using of computerized-matchmaking called collaboration filtering. Finally, two experimental results demonstrate that the personalized algorithm can produce a better predication than the non-personalized one and that system performance improves when the number of user rating increases.

IM-03-14

DESIGN PATTERNS FOR MOBILE APPLICATIONS DEVELOPMENT

Ms. Prodpran Muensuksaeng

Dr. Dencho N. Batanov

Due to the growth of mobile world, the rapid change of technology in today's networked world presents a great challenge for system developers in term of shorten time of development. Implementing every system from scratch does not scale up to the demand. Therefore, the combination of object-oriented programming with pre-build software components would satisfy the development strategy. One of the pre-built software components is design patterns. This study proposes that the existing patterns must be modified to solve the limitations of mobile devices and the new patterns need to be developed. The study starts up with analyzing of mobile application development and their constraints. Then this study analyses the chosen existing design patterns. Subsequently, the chosen design patterns are modified and also creating of new ones to suit mobile application limitations and constraints are examined. Ultimately, the study comes up with new design patterns, which are Mobile adapter, Synchronization, and Caching patterns. Finally, a prototype of home library system is developed as an illustrative case to show employing of the new design patterns combining with technologies of object-oriented, UML, WAP, and design patterns together in order to achieve the study.

IM-03-15

A FRAMEWORK FOR BENCHMARKING A TRAVEL WEB SITE

Ms. Pwint Phyu Aung

Dr. Vatcharaporn Esichaikul

With the advent of the Internet, travel web site businesses are rapidly growing. Consequently, it is important to determine the performance of the web site. In most industries, business performances can be compared with one another by applying business frameworks. However, there is no proper framework for evaluating travel web sites. This study is an attempt to propose criteria for benchmarking a travel web site. The result of this study provides such a framework. A survey was conducted to obtain information from travel web site owners or developers and travel web site users. The results of the survey revealed important criteria for the study. These criteria were differentiated into general criteria and technical criteria. Analysis was conducted by using a statistical approach and software tools. The analysis results were used to implement the prototype. Finally, the study proposed a framework to determine benchmarking criteria for travel web sites. This framework would be useful for the travel web site developers or owners who want to compare their web site with others for purpose of improvement. Travel web site users can also use this benchmarking framework to compare travel web sites to identify ones that meet the criteria.

IM-03-16

WIRELESS AND BLUETOOTH TECHNOLOGIES FOR ENABLING WEB APPLICATIONS

Ms. Morakot Apinyankul

Dr. Dencho N. Batanov

Currently, the need of information that is available on the Internet is growing very fast, particularly in Wireless communication industry. Users also require information for serving anywhere, anytime, and any devices. Therefore, the feature's keys of Web Applications are not only serving information but also providing high available, reliable infrastructure, simplest, easiest, and cheapest to the users. Regarding these requirements, Web Application on Wireless device architectures has been proposed for fulfilling user's demand by using distributing computing (Client-Server computing) and Bluetooth Internet Bridge. Applying Wireless and Bluetooth Technologies for enabling Web Applications is designed by four main objectives that are simplifying Web Applications, reducing cost of browsing Web Applications, enabling ease of use Web Applications, and improving Web applications performance. To achieve these objectives, the following services are proposed and implemented in the prototype system that are MIDlet J2ME client-side application, server-side application on Web Server Environment, and Bluetooth Internet Bridge.

This system is designed for providing Web Applications in various client devices and each device needs an individual technology supported, which is usually different from others. However, the system should allow shared Web Application by applying the same business logic to all devices within the entire system. Therefore, this Web Application framework supports Wireless devices (3G/Bluetooth) and portable computers.

IM-03-17

TRANSITION FROM ASP-BASED LEGACY SYSTEMS TO .NET-BASED SYSTEMS

Mr. Shezad Farid

Dr. Dencho N. Batanov

Microsoft .NET is the latest generation of software technologies that makes it easier to share information between computer systems. It allows web development giving the developer a notion similar that of developing a desktop application. It is a generation ahead of its predecessor ASP in terms of architecture and performance. ASP.NET has tried to work on many of the shortcomings of ASP and added features like object-orientation and multiple language support; it has also evolved into a compiled module. To take advantage of this new technology one can transition existing ASP applications through a prescribed strategy and a set of guidelines. The resulting guidelines aim to provide a feasible transition to the new environment with minimum amount of workload, such that maximum amount of resources could be utilized from the present system to work on the new environment. The prescribed strategy and a set of guidelines have been implemented on the Q&A e-learning system.