

ABSTRACTS-YEAR 2002

THESIS

COMPUTER SCIENCE

CS-02-01

A CASE STUDY IN THE USE OF KNOWLEDGE-BASED ARTIFICIAL NEURAL NETWORKS FOR PRODUCT SELECTION DECISION SUPPORT

W.P.C. Duminda de Silva

Dr. Peter Haddawy

In this study, we explore the use of Knowledge Based Artificial Neural Networks (KBANN), pioneered by Shavlik and Towell, 1994 [25], to learn user preferences under certainty. We start by describing the problem of choosing a used motorcar, where it is reasonable to make several assumptions about preferential independence and monotonicity. We then show how to represent these assumptions as Horn-Clause theories that can be encoded in a KBANN. We then empirically compare the KBANN with simple back propagation Artificial Neural Network (ANN) in terms of learning rate and accuracy. We also examine the performance of KBANN in learning preferences using examples generated from a number of value functions that violate the assumptions to various degrees and compare the results with that of ANN.

CS-02-02

AN ANALYTICAL STUDY OF PLANNING

Dehigaspitiyage Don Ananda Gamini

Prof. Phan Minh Dung

Over the years, forward chaining planners have fallen out of favor in the AI planning community, and partial order planners have shown to possess a number of advantages. The experimental results showed that forward chaining planners, by utilizing additional information in the domain, could scale up to very larger problems in many domains. The use of efficient domain information could make intractable planning problems tractable; in particular, SHOP (Simple Hierarchical Ordered Planner) could be able to reconfigure 500 blocks in about 26 minutes. SHOP outperformed TLPlan in most of the domains under study. The primary reasons for the differences in the performance of two planners are: (1) TLPlan's planning algorithm is basically a state-space search, whereas SHOP uses HTN-style problem reduction, and it has been established that the problem reduction can be much more efficient than state-space search, and (2) Even though TLPlan's modal-logic representation capabilities are quite sophisticated, their use has been limited for writing pruning heuristics rather than actual planning algorithms.

CS-02-03

DESIGNING SINHALA CHARACTERS USING BÉZIER CURVES AND B-SPLINE CURVES

Thantulage Gishantha Indunil Fernando

Prof. Huynh Ngoc Phien

In character design in the computer industry, the mathematical models like Bézier curves and B-Spline curves play a very important role. Earlier characters were stored in the bit maps, and the designer had to keep separate file for each size of each font. But after Adobe introduced the PostScript fonts into the market, the fonts were represented by Bézier curves and those fonts can be scalable to any size. Further, this needs only one or two files for storing the entire font type.

Existing methods of representing characters and theoretical background about Bézier curves and B-Spline curves are reviewed in this study.

Further, we have developed a Visual Basic program called "Fontica" for designing characters by using Bezier and B-Spline curves. With the help of this new program we could design 10 Sinhala (my mother language) characters.

In addition, we have proposed a new technique for approximating the degree reduction of Bezier curves. This approximation appears to perform satisfactory.

CS-02-04

METADATA-DRIVEN DATA WAREHOUSING WITH XDD

Paramai Supadulchai

Prof. Vilas Wuwongse

OMG's Common Warehouse Metamodel (CWM) is a standard modeling language for designing data warehouse systems. Nevertheless, it provides mainly syntactic metadata for either technical or business domain while it lacks metadata reasoning approach to enable metadata computation required by modern data warehouses. This study proposes a framework to represent and reason with semantic metadata in order to capture major data warehouse behaviors. These behaviors can be implicitly employed to manipulate syntactic metadata in certain circumstances. This framework seamlessly works for CWM compliance tools as well as proprietary ones mostly found in present data warehouse solutions where tools cannot be selected from a single vendor.

By employing XML Declarative Description (XDD) as data representation and Equivalent Transformation (ET) as inference mechanism, the proposed framework can process CWM models represented in XML Metadata Interchange (XMI) format and

derive new models resulting in appropriate configurations and dynamic system behaviors. In order to demonstrate the effectiveness of the proposed framework, a prototype system has been developed. This system can be applied to various data warehouse behaviors. In this study, development of data warehousing using prototype approach is selected as a case study in order to show the mechanism of the framework.

CS-02-05

SEMANTIC ANNOTATION OF MULTIMEDIA

Ano Chotmanee

Prof. Vilas Wuwongse

MPEG-7 provides a set of tools for the definition of a standard description for multimedia information. It uses XML Schema as its own Description Definition Language (DDL) to provide an ability to extend standard description to user application domains. However, due to its lack of expressive power, XML Schema cannot represent the implicit information of multimedia contents and hence cannot yield derived descriptions. In order to enhance the expressive power of multimedia annotations—described in term of an XML document—and to maintain the advantages of XML Schema, there appears to be a need of a new schema language, which can handle those abilities. XML Declarative Description (XDD) is a representation framework, which can encode any XML document and application (e.g., XML Schema, RDF, RDF Schema, and DAML+OIL). It enhances XML expressive power by employing Declarative Description (DD) theory and it also provides a mechanism for reasoning about new information from existing data. Moreover, it can specify constraints and modeling ontological axioms.

CS-02-06

MODELING AND REPRESENTATION OF KNOWLEDGE FOR KNOWLEDGE MANAGEMENT

Penpaka Rongmarood

Prof. Vilas Wuwongse

Knowledge representation frameworks are a fundamental part of knowledge management systems. Normally, a framework consists of three interrelated components: knowledge modeling technique, knowledge representation language and reasoning mechanism (Inference Engine). Many frameworks have been developed with the common goal to provide sufficient expressive power to model and represent all kinds of knowledge in a clearly specified syntax, formal and precise semantics as well as efficient reasoning support. Moreover, a knowledge representation language has to be formulated by means of existing web standards for interchangeability among applications. However, at present there is no framework that can fulfill this goal. Therefore, a new knowledge representation framework is proposed aiming to accomplish that goal. Its design is based on three concepts: standard unified knowledge modeling language by means of Unified Modeling Language (UML), unified knowledge representation language by means of XML Declarative Description (XDD) and efficient reasoning mechanism by means of Equivalent Transformation (ET). These three concepts lead to the provision of a practical and efficient inter-knowledge representation framework. Based on this framework, a prototype of knowledge management system for a research organization is developed as a demonstration of its practicality and potential usage in real-world knowledge management.

CS-02-07

XML SCHEMA WITH ENHANCED SEMANTIC CONSTRAINTS AND ITS MODELING

Patcharee Thongtra

Prof. Vilas Wuwongse

XML Schema is an XML schema language used mainly to describe the content model of an XML element. A content model usually specifies information about sub-elements and their relationships or structure. However, in real-world XML documents there are others kinds of information or relationships that need to be specified, e.g., the content value of an element is equal to the summation of those of some other elements. This kind of relationship is called a semantic constraint. This study first proposes a framework that can represent such semantic constraints by enhancing XML Schema. A technique is then developed to model an XML Schema with semantic constraints and validate an XML document against it. Finally, an electronic-form submission system with semantic constraint validation is designed and implemented in order to demonstrate.

CS-02-08

A COMPARISON OF FULL AND PARTIAL RECURRENT NEURAL NETWORKS

Rattaphon Sakchaicharoenkul

Prof. Huynh Ngoc Phien

Recurrent Neural Networks (RNNs) are one in which self-loops and backward weight connections between neurons are allowed. As a result of these network characteristics, recurrent networks can address temporal behaviors which not possible in Feed Forward Neural Networks (FNNs), such as their behavior in the limit reaches a steady state (fixed point), an oscillation (limit cycle), and an aperiodic instability (chaos). Since RNNs have been increasingly applied to many dynamic system applications, there have been extensive efforts to develop a variety of architectures and training algorithms concerning on the enhancement of dynamic system characteristics.

This work focuses on comparison of selected architectures between full and partial RNNs and also compares selected training algorithms, namely Error Back Propagation and Exponentially Weighted Least Squares algorithm, Accelerating Convergence Using Approximated Gradient algorithm, and Error Self-Recurrent Back Propagation with Recursive Least Squares algorithm. Empirical comparison is also made with respect to time and accuracy using three data sets: daily stream flow (rainfall-runoff) data, quarterly data on exports and gross domestic product (GDP) of Thailand, and daily data on stock prices in Thai market.

A proposed algorithm was obtained by applying the estimation of pre-image signals at the hidden layer with the fastest training

algorithm among several algorithms considered. The algorithm was devised to speed up the convergence of networks and to enhance the network performance considered with the comparative algorithms. It was found that the proposed algorithm performed much better than the selected training algorithms when it was applied with Partially Recurrent Neural Networks (PRNNs).

CS-02-09

A CONNECTIONIST MODEL OF ENGLISH-TO-VIETNAMESE MACHINE TRANSLATION

Nguyen Hong Chuong

Prof. Huynh Ngoc Phien

This study concerns with investigating and building a connectionist model of English-to-Vietnamese machine translation. The proposed model integrates all the connectionist tasks into a single and coherent framework. In addition, this model could be developed into an interlingua system that can translate an English sentence into many sentences of many natural languages (one-to-many interlingua system).

The SEnglish and SVietnamese languages are the two subsets of English and Vietnamese. They are used as the source and target languages of the proposed model. SEnglish defines the relatively complex sentences by a stochastic context-free grammar. The meaning of a SEnglish sentence is represented by a set of propositions. SVietnamese is only defined by a context-free grammar. A SVietnamese sentence is generated from a set of propositions.

In this model, a recurrent neural network is trained to understand a relatively complex SEnglish sentence by translating it into a set of propositions, which will be translated into the SVietnamese sentence later.

This study also implements and tests the proposed model by a very simple SEnglish language whose sentences have only one proposition.

CS-02-10

HYBRID IMAGE REPRESENTATION USING QUADTREE AND OBDD

Handoko

Dr. Kanchana Kanchanasut

Ordered Binary Decision Diagram (OBDD) has been used to reduce the amount of space and computational required for verifying digital circuits by removing copies of sub-functions in Boolean function. An image can be processed as Boolean function as it is assigned with variables along x and y coordinates. The number of variables in OBDD will effect to the performance of OBDD in terms of ratio and time. Quadtree is a spatial image compression algorithm that has been widely used to represent images such in GIS or medical X-Ray. It is a fast algorithm that divides an area into quadrants and recursively divides each quadrant into smaller ones until the area cannot be divided into smaller parts anymore. By combining these two algorithms, an image can be processed using various sizes of OBDD depends on the characteristics of block that will be processed.

Lossless image compression using Quadtree-OBDD typically produces average compression ratios 1.25 for natural images, 3.29 for text images, 7.49 for binary images, 30.19 for synthetic images and 1.02 for texture images. These results are better than GIF algorithm for all image classifications except text images and better than Lossless JPEG for binary images.

CS-02-11

DESIGN AND IMPLEMENTATION OF A RECONFIGURABLE COMPUTING SYSTEM FOR DIGITAL IMAGE PROCESSING

Ramiro Ivanez Catala

Dr. Lertsak Lekawat

An important part of the communication and data processing industry needs the development of prototypes for testing designs. This master thesis is fully oriented to practical issues of digital image processing system design and the target is to design and implement one of these prototypes based in Field Programmable Gate Array and microcontroller. The intention was to make the design in a concrete framework and hence a real world application has been studied: geometric correction based in polynomial approximation.

This project belongs to an exchange program between the Asian Institute of Technology (Bangkok) and the University of Applied Sciences of Landshut (Germany), the former is the institution will evaluate this research as a thesis for a Master Degree in Computer Science, the later is in contact with several companies of the electronic design sector in Germany and it could provide a professional environment for the development of this project. Specifically, this thesis has being carried out in Isartec, located in Landshut, under the supervision of Mr. Markus Waidelich (Isartec) and Dr. Fritz Poernbacher (fh-Landshut).

CS-02-12

DIRECTORY-ENABLED NETWORK BASED INTELLIGENT NETWORK MANAGEMENT SYSTEM

Dao Hoang Duong

Prof. Phan Minh Dung

The thesis aims at developing an application with a graphical user interface to manage network system based on DEN specification applied in ISO management of Fault and Configuration. The application database has been developed in hierarchy and stored in LDAP. Therefore, it is very easy and convenient for the network administrator to have a look from the

overall level down to the lowest one on his network as well as separate elements. Also the application integrates with SNMP agents to provide the ability of obtaining and updating information based on the user's opinion. The system is constructed in one of the latest programming language of Java SDK with the support of Forte for Java. Any modification in the system will be performed very easily whenever needed. The same methodology of the study can be applied to further improve the application before implementing in the real life.

CS-02-13

LOGIC PROGRAMMING BASED PROOF PROCEDURE FOR TRUST MANAGEMENT

Nguyen Duy Hung

Prof. Phan Minh Dung

Computer security needs a language to express security policies and a protection mechanism to control users' access to protected resources according to given security policy. Hence, at the core of a protection mechanism is a compliance-checking algorithm, which checks whether a user's request complies with the policies. In distributed systems, Trust Management (TM) is the approach towards computer security that groups together several principles such as authorization information is in formed of digitally signed credentials and local policies expressed by an application-independent language. Role-based Trust management (RT), a recently proposed framework is claimed to have combined strengths of previous TM systems. Li and Mitchell and Winsborough (2001) have designed direct graph-based compliance checking algorithms for RT0, which is the simplest language of RT framework. Algorithms for other advanced languages of RT framework are still open.

In this thesis, we build compliance-checking algorithms based on logic query evaluations. The results adopt that logic query evaluations can be optimized so that they are reduced to direct graph-based algorithms. Hence, the latter can be seen as specializations of the former. We further conclude that graph-based representation is not new at all as it is claimed to be. We believe that the logic-based approach offer advantages: (1) we don't have reinvent the wheel ;(2) we can generalize when the language involves.

CS-02-14

TRAINING ALGORITHMS FOR RECURRENT NEURAL NETWORKS

Nguyen Nhi Gia Vinh

Prof. Huynh Ngoc Phien

In recent years, artificial neural networks have been applied in many fields, especially recurrent neural networks have attracted a lot of attention and important focus in research and publication. A recurrent neural network is one in which self-loops and backward connections between nodes are allowed. One of the consequences of this fact is that their dynamic behaviors not possible with strictly feed forward neural networks, such as limit cycle and chaos, can be produced with recurrent networks. The diversity of dynamic behavior leads to the situation where recurrent neural networks well suit to many important problems including filtering and forecasting. Another possible benefit of recurrent neural networks is that smaller networks may provide the functionality of much larger feed forward neural networks.

Despite the potentiality and capability of recurrent networks, the main problem is the difficulty of training them: the complexity and slow convergence of the existing training algorithms. This study focuses on understanding recurrent neural networks dynamic behaviors and on improvements of algorithms to achieve faster convergence. An experiment is carried out by means of the use of two sets of rainfall-discharge data for the fully recurrent networks and Elman (partially) recurrent networks.

The results obtained show that:

- Autocorrelation and cross correlation analysis can be used to determine the number of input nodes
- The number of hidden nodes can be determined by combining the Baum-Haussler rule and Bayesian Information criterion as proposed in this study
- In terms of the computation time, the offline training algorithm of Atiya and Parlos is fastest, while the real time learning algorithm is slowest
- In terms of the performance statistics, the offline training algorithm of Atiya and Parlos has the highest performance.

CS-02-15

A STRAND-BASED LOGIC PROGRAMMING APPROACH TO ANALYSIS OF AUTHENTICATION PROTOCOLS

Nguyen Viet Anh

Prof. Phan Minh Dung

Authentication protocols play an important role in any distributed system, especially in the Internet. However, it is surprisingly strange that most protocols do not achieve their goals despite the fact that they are carefully designed and examined. Hence, it is of great important to develop a formal method for checking the correctness of those protocols.

So far, many such methods have been proposed. Of which, strand space model (SSM) is a simple and elegant way for reasoning about authentication protocols. The disadvantage of SSM is that the proofs are complex and in ad-hoc manner. In this thesis, we propose a strand based logic programming for automating the process of analyzing authentication protocols. The logic programs are executable via the general models.

CS-02-16

FIREWALL CONFIGURATION FOR PRIVACY

Phan Van Tri

Prof. Phan Minh Dung

The Internet is a global information world containing valuable knowledge and information. In its public spaces, people interact commercially and socially with strangers as well as with acquaintances and friends. Thus, in order to explore and use efficiently these information resources, we must filter to select the useful information and also protect our resources from people with malicious intent. Firewall is one of the most effective tools to protect a network from outside attackers.

Firewall is a system designed to prevent unauthorized access to or from a private network. Firewalls can be implemented in both hardware and software, or a combination of both.

In this report, I surveyed the current status of the real network of Electricity of Vietnam (EVN), and researched to propose a network architecture that is most suitable for upgrading the WAN network of EVN. I also design a firewall to protect the resources of the EVN network and to filter the inappropriate materials or materials that are not related to the job of the employees. The firewall must flexibly satisfy the diversity of policies and it must be easy to reconfigure when the policies are changed. The Firewall must also dynamically filter web sites that are newly encountered. The firewall solution that is proposed to filter Internet materials will be built based on the Platform for Internet Contents Selection (PICS).

CS-02-17

HIERARCHICAL CLUSTERING OF PRODUCT DATABASE FOR FAST UTILITY-DIRECTED SEARCH

Tran Van Dung

Dr. Peter Haddawy

Utility or value function can be used to represent preferences. In many applications, there is a need to search a database of products for a product or set of products that have highest value computed from a submitted utility function. Such application may be implemented with the help of a database management system. However, with the relational database management system, we need to compute the value of every product in the database; therefore time complexity will be large $O(N)$. This is acceptable with small database but it will significantly slow down the overall performance of the system with large database.

This study is to apply hierarchical clustering to the problem of product database management to improve searching speed. The author discusses a technique for building the database, which uses product abstraction in hierarchical form to speed up the utility-directed searching speed. The efficiency of the proposed technique is evaluated by comparing its performance to that of a system, which relies on a relational database management system.

Key words: Hierarchy, cluster, utility function, preference structure, indexing.

CS-02-18

NETWORK MONITORING

Nguyen Canh Toan

Prof. Phan Minh Dung

Today's enterprise networks are composed of multiple types of interconnected networks. On top of these enterprise networks, there exist various systems and services supporting a wide variety of applications within an organization. Providing a secure, reliable and efficient operating environment to support the organization's daily activities and its business is the most challenging task faced by operation and management staff today.

In order to provide such an environment, enterprise networks must be monitored for performance, configuration, security, accounting and fault management. Current management practices typically involve the use of complex, hard-to-learn and hard-to-use tools for managing networks. What is needed desperately is a set of simple, uniform, ubiquitous tools for managing networks. Web-based management is a promising approach that can provide such a solution.

This research focuses on the use of Web technology for the purposes of enterprise network traffic monitoring, analysis and reporting. In this research, we first examine the requirements for enterprise network traffic monitoring, analysis and reporting, and then present a design and implementation of a Web-based network traffic monitoring and reporting system that satisfies those requirements.

CS-02-19

SELF-STABILIZING MULTI-AGENT SYSTEM FOR DISTRIBUTED TREE GRAPHS

Nguyen Thi Phi Loan

Prof. Phan Minh Dung

Nowadays, the distributed systems are applied more and more and become the indispensable needs in the real life. However, due to the susceptible-to-failure nature of these systems, the self-stabilizing algorithms are proposed as a way to handle failures and bring back the system to the legitimate state.

A number of self-stabilizing algorithms are proposed in many different problems such as spanning tree, leader election, minimum spanning tree and so on. However, almost of these algorithms are written by the lower level language. The complexity of the algorithm itself and the limit of language in expressing cause the algorithms more difficult to understand.

This thesis classifies some problems of the self-stabilizing algorithms and transforms them into the multi-agent systems (MAS).

CS-02-20**MEASURING COUPLING AND COHESION FOR HIGH QUALITY OBJECT-ORIENTED SOFTWARE DEVELOPMENT****K.M. Azharul Hasan****Dr. Dencho N. Batanov**

Cohesion and coupling are internal software attributes, which tell how tightly the components of software module are bound together in design or implementation. Cohesion and coupling are the desirable goal in software construction, leading to better values for external attributes such as maintainability, reusability, and reliability.

Object-oriented (OO) system development is gaining wide attention both in research environment and in industry. A severe problem is encountered, however, is the quickly increasing complexity of such systems and the lack of adequate criteria and guidelines for good designs. To cope with this problem it is imperative to better understand the properties and characteristics of object-oriented systems. Cohesion describes the binding of elements within one method and within one object class, respectively. Coupling describes the interdependency between methods and between object classes, respectively.

Considering all the features of object-orientation the coupling in object-oriented system has been categorized. A model has been presented for measuring direct coupling of OO system, which is then used to identify components with the objective of encapsulating higher coupling objects into the components. Some ways has been suggested for which an OO approach can be developed with high cohesion and low coupling. Relationships between two well-known and widely accepted measures of cohesion have been shown. The effects of Public members, Private members, Constructors, Destructors and Inheritance in measuring cohesion have been analyzed through principal component analysis. Some statistical analysis has been shown to validate the proposed measures.

CS-02-21**PREDICTING SUPPLY PURCHASE PATTERNS IN DATA-RICH E-MARKETPLACE ENVIRONMENTS****Yongyos Kaewpitakkun****Dr. Peter Haddawy**

Barter system is one kind of trade exchange. Through barter, people can get things that they want without paying money. They exchange their products or services instead. With barter, the medium of this commerce is called "trade dollar." People can earn trade dollars when they sell their products or services into the product pool. And they can use their trade dollar to buy any thing in the pool.

With the high growing rate of barter business, there are many barter trade exchanges. They can earn commission on every transaction occurred through them. But, surprisingly, there is no complete software for automating the barter process. Most of transactions still made through human broker. This situation make the barter trade exchange can not reach their maximum profits. Therefore, this thesis is concerned with generating the automated potential recommendation for support in barter business.

In this thesis, techniques from machine learning, such as Bayesian network, and collaborative filtering are integrated to develop a candidate recommendation engine for predicting supply purchase in data-rich e-marketplace environment. We have considered several continuations, grouped into 4 categories: generated the recommendation based on a single company itself, generated the recommendation based on a group of similar company, validation the result, and improvements to the ordering search space for K2.

CS-02-22**RETRIEVAL AND COMPOSITION OF TEMPORAL AND SPATIAL MULTIMEDIA****Napat Sukthong****Prof. Vilas Wuwongse**

This research first proposes a general framework for retrieval and composition of temporal and spatial multimedia. This framework is based on MPEG-7 and SMIL standard those provide the tools to annotate the temporal and spatial relations inside multimedia and the tools to compose the multimedia, respectively. SMIL is selected to be one of the representation languages for the multimedia documents and presentations but it has not enough efficiency to handle the higher-level temporal and spatial relations. MPEG-7 provides a tool to annotate these kinds of relation e.g. multimedia segment decomposition tools, segment relation description tools, temporal segment relation description tools, spatial segment relation tools and spatio-temporal segment relation description tools. These tools used to decompose multimedia contents into segments and synchronize them together. Nevertheless, they should have more expressiveness power for retrieval and composition. Thus, the XDD theory is employed to enhance the language become expressible and computable. Furthermore, by employing XDD, constraints and axioms specifications are easily defined. This paper also defines a modeling language called SMILD that be able to represent temporal and spatial multimedia.

CS-02-23**WEB SERVICES AND ENTERPRISE APPLICATION INTEGRATION****Winai Wongthai****Prof. Vilas Wuwongse**

Web Services enable machine-to-machine interoperation in a distributed computation system. Deployment of their full potential by Enterprise Application Integration (EAI) is the main aim. The present methodology provides an appropriate conceptual model for the development of Web services from legacy applications. Initially, the legacy environment approach and building of the connector to a certain legacy environment by means of Java is focused on demonstration. Subsequently, at the

implementation phase, the communication between Java connectors and legacy files and databases is demonstrated. A prototype tool is then applied for generating certain necessary files from a Java connector in several wrapping scenarios as Web services and invocation of service from client applications is described.

Various technologies, currently available, are also referred to as well as their employment in the development of Web Services, EAI, Java Technology and Web service tools.

CS-02-24

A LOGIC PROGRAMMING-BASED FORMALISM FOR LOCAL CLOSED WORLD REASONING IN PLANNING

Dinh Trung Hang

Prof. Phan Minh Dung

LCW mechanism is one way to handle incomplete information in planning, which was first investigated by Etzioni et al. and later advanced by Bacchus and Petrick. This work is another approach to LCW information in planning based on logic programming. We describe how LCW information is represented and inferred under logic programming framework. We also provide a high-level action language A_L to capture all properties of actions that involve LCW information. The focus of defining this language is on semantics of knowledge updating at planning time and at execution time. Our formalism is shown to be sound with respect to reasoning about knowledge and actions.

CS-02-25

QUERY PROCESSING OF INTERGRATED XML DATABASES

Le Thi Thu Thuy

Prof. Vilas Wuwongse

Query processing of integrated XML databases has increasingly gain attention in the last few years. It helps users to extract information from an integrated XML databases system. The big challenge of query processing of integrated XML databases is to find a suitable model for query processing, metadata representation and data instance integration. Such a model will increase the flexibility of query processing processes as well as serve users with friendly integrated data.

The proposed framework displays any successes. Through a single request, without concern about their internal structures and geographical distribution, it allows decomposition of global queries into sub-queries conformed simultaneously to local format. In order to reduce the redundancies of mappings, low-level metadata processing has been introduced; it yields two direction-mapping for query decomposition and data conversion from the original metadata. It is also shown how to extract data to provide user-friendly results. XML Declarative Description is applied to all components of the framework as their underlying model.

CS-02-26

AN ANALYTICAL STUDY OF BUFFER OVERFLOW VULNERABILITIES

Tran Thi Minh Chau

Prof. Phan Minh Dung

Over the last decade, buffer overflow have been the most common form of security vulnerabilities. Buffer overflow attacks are one of the most serious classes of security threats, as they enable attackers to gain total control of a host computer locally or remotely.

This report is the result of a study to understand the principles of Unix/Linux buffer overflow vulnerabilities, model the essence of buffer overflow. Next, the study surveys existing static detection approaches, and works towards a better method of buffer overflow static detection.

CS-02-27

INCREASING USER FLEXIBILITY IN CREATING AND CHOOSING CHANNELS ON PUBLIC PORTALS

Bui Quoc Thai

Dr. Peter Haddawy

Navigating and managing information on the web for each person is an important topic today. With the explosive growth of the Internet and World Wide Web, users are overwhelmed with information. A method that can help the users retrieve information effectively needs to be provided.

A Portal is a mechanism that enables people to navigate and collect information effectively. It is like a "door way" or a "single entry" at which users can drop by and see different things on the web. In other words, portals can collect data from different sources on the web and present them on a single page in an individualized manner. Moreover, portal users can customize their own page with interesting color, layout, and contents.

In this study, an approach for increasing user flexibility in choosing channels is researched. My study will help web surfers overcome the restriction of most current portals – portals just provide a fixed set of channels, and hence limit the content that users can view. That means the study gives a solution that allows users to create channels for themselves. This not only increases the flexibility for the users but also eliminates the tedious task in creating channels from portal administrators. Besides, a recommendation method is developed to predict and recommend to the users what information channels the users may like. Through this recommendation, the users can discover or create the channels for themselves.

CS-02-28

XML DATABASE SCHEMA INTEGRATION

Doan Dai Duong

Prof. Vilas Wuwongse

XML databases schema integration has lately received increasing attention. It helps to obtain unified representation of all participating databases of an organization, facilitating their information access and utilization. The presented XML database schema integration framework follows one short strategy. The system is capable of integrating simultaneously n schemas. Possible conflicts between XML schemas are detected and their resolutions proposed. A normal form of XML schema is also given; it provides a unique declaration of an XML schema and avoids ambiguous schema representation. XML Declarative Description is used as the framework's underlying model.

CS-02-29

TRANSFORMATION OF XQUERY INTO XML DECLARATIVE DESCRIPTION (XDD)

Nguyen Manh Cuong

Prof. Vilas Wuwongse

Increasing amounts of information are stored, exchanged, and presented by means of XML. In order to achieve effective manipulation of a large collection XML documents, there arises an essential need for an XML standard language, which allows users to query, retrieve, transform and construct XML data which meet precise requirements. Xquery -an XML language query -provides features for retrieval and interpretation of information from diverse XML document sources.

However, so far, there is no unification with respect to the standard XML query language. Although, Xquery has been recommended by W3C, it is not yet the latest version; it is changing continuously month by month. Hence it is not easy to build an engine for Xquery's execution of a query. Hitherto, Microsoft has developed Microsoft's XQuery Demo, which, however, is only a demonstration version and has many drawbacks.

This is an approach to XML query formulation and evaluation developed by means of XML Declarative Description (XDD) theory. It comprises one or more XML clauses so that its syntax can be subdivided into three specifications: Pattern of XML elements, criteria and the structure of the resulting elements. It supports formulation of essential functionality requirements for an XML query language such as selection, extraction, combination, transformation for specified XML databases.

CS-02-30

REASONING WITH UML STATECHART DIAGRAMS USING XML COMMUNICATION PROTOCOLS DECLARATIVE DESCRIPTION THEORY

Nguyen Tri Thanh

Prof. Vilas Wuwongse

The Unified Modeling Language (UML) is a de facto standard language for analyzing and designing system in object-oriented approach. One of its important types of diagrams, statechart diagrams are complementary views of many other kinds of diagrams in UML. It helps describe dynamic behaviors of its attached diagram. This study proposes a representation the semantics of UML statechart diagram using XML Declarative Description (XDD). Other objectives are to find out rules to verify the correctness of the input UML statechart diagrams in XML-based format, refine these diagrams, checking consistencies among statechart diagrams and its attached diagram, transform them into other implicitly equivalent diagrams.

XDD is used as a knowledge representation tool. Graphical UML diagrams are converted into XML-based format and become facts in a knowledge base. All rules for verifying, refining, checking, transforming are represented as a set of XDD rules in XML format. The XML Equivalent Transformation (XET), a programming language that is based on Equivalent Transformation (ET), is used to process UML statechart diagram encoded in XML format.

CS-02-31

A DIGITAL LIBRARY SYSTEM OF THESESES

Phan Vo Minh Thang

Prof. Vilas Wuwongse

A digital library system is developed to manage theses and dissertations. It employs open-source components and can store a large volume of theses, organize theses into collections and export these collections to CD-ROM. This system allows search by contents and has a user-friendly interface. It also employs Dublin Core elements in order to enable search by metadata.

CS-02-32

SELF-STABILIZING MULTI-AGENT SYSTEM FOR COMMUNICATION PROTOCOLS

Nhuyen Thi Tuyet Le

Prof. Phan Minh Dung

An important function of communication networks is to implement reliable data transfer over an unreliable underlying network. Research in self-stabilizing aims at making distributed systems more reliable by handling faults in complex computing environment. A communication protocol is self-stabilizing if it guarantees that the protocol will eventually reach a legitimate state when started from an arbitrary initial state within a finite number of state transitions. Stabilization allows the processes in a protocol to re-establish coordination between one another, whenever coordination is lost due to some failure.

The aim of this thesis is to study general techniques to make communication protocols self-stabilizing. We also discuss how to apply these techniques to transform the sliding-window and the two-way handshake protocol to make them self-stabilizing.

Finally, we convert the protocols into multi-agent systems using the Agent Programming Language, which is a high-level language, which variables come with predicates that define them.

INFORMATION MANAGEMENT

IM-02-01

DEVELOPMENT OF SMART ACTIVE MOBILE INTERNET SERVICES

Jukka P. Silvennoinen

Dr. Dencho N. Batanov

We are living in the edge of new dawn for mobile Internet services. The so-called 3G Mobile technologies are starting to emerge into the market. Newly introduced technologies in hardware, protocols and operating systems in past years are giving mobile Internet new possibilities to serve the customers. The introduction of technologies and new communication methods has also its down side. Increased divergence in client capabilities makes the design of the services more complex issue than what it used to be before these technological introductions.

One other trend in mobile market is the falling revenue of the traditional voice communications. The data communication is seen as a method of making the mobile business more profitable and to overcome the problem of falling revenue. This means that the service application should be build to serve as many communication protocols and device families as possible to maximize the usability as well as the profitability of the mobile Internet service.

The problem is that there is no universal method to serve all possible communication methods and also same time delivers the content in a way that it suits the client capabilities best.

To solve the problem as a result of this study a mobile Internet service application structure is introduced. The structure can be built and modified to suit a wide variety of different communication methods and also it has capabilities to modify the content to suit client's capabilities as needed.

Also the structure doesn't address the implementation issues, so it allows the implementation to be done using the methods that will suit the solution provider's interests best.

This study introduces two prototype applications that illustrate the usability and the flexibility of the proposed structure with real word examples.

IM-02-02

SERVER CLUSTERS AND SERVER FARMS: A CASE STUDY IN AIT NETWORK ENVIRONMENT

Sari Retulainen

Dr. Vatcharaporn Esichaikul

This thesis takes an example of the modern type of server cluster and weighs the pros and cons of acquisition of such implementation. The basic information to understand such arrangements has been gathered from the Internet sources. The information providing views onto the case study environment have been gathered from personal interviews of the administrative personnel of this specific environment. Part of the information comes from the documents concerning the environment management and development.

When joining the information from these separate sources together one can see some discrepancy between the theory and practice in the case study environment. The implementations seem to be deficient in some parts and the strategic planning seems to be too cursory. The detailed planning considering the whole system is missing.

Therefore this thesis illustrates how clustering solution could improve the whole system performance by providing better facilities for some parts of the system and thus taking the extra load off the other parts of the system.

IM-02-03

MUSIC XML COMPUTATION FRAMEWORK AND ITS APPLICATIONS

Sarawadee Ekpanayapong

Prof. Vilas Wuwongse

With growing music publication over the Internet, there has arisen an increasing demand for a music representation and computation tool. Many efforts have been made to compute music, but most of them do not permit direct processing of music documents, because they lack expressive tools for direct processing and their computational tools are limited to description of musical knowledge. A MusicXML computation framework is proposed both for music representation and computation. In it, music representation and computation are done by a unified modeling language *XML Declarative Description (XDD)*. *XDD* partially employs MusicXML as a format for music representation. With this framework, music representation and computation occur in a uniform representation of music notation and musical knowledge. Direct computation of music documents and flexibility of knowledge modification characterize this framework.

IM-02-04**PERSONALIZED META-SEARCH ENGINE**

Patipat Korbsook

Dr. Peter Haddawy

This thesis describes a method for a personalized meta-search engine. The proposed method produced a personalized search result by combining the concept of explicit and implicit interest of user. The user interest and preferences are stored as user profile. The profile is then used as the result ordering policy. Rocchio algorithm is employed to adapt the term weights in the user profile so that user's preferences are best represented. The system divided searching into two type of searching, short-term for the short-term interest search and long-term for the long-term interest search. Both type of searches used the user profile as the results ordering policy. The different is that, the long-term search method analyzes user's reaction to the represented documents. The user action is directly observed and learned to extract the user interest implicitly. The proposed method show superior performance in personalized search result in both type of searches compared to the result ordered by conventional meta-search engine.

IM-02-05**METHODOLOGY FOR DEVELOPMENT OF WEB APPLICATIONS. CASE STUDY: USING ASP.NET FOR INTERNET BANKING**

Panu Leelaluckana

Dr. Dencho N. Batanov

In the world of Web application development, ASP.NET is the latest Microsoft technology that provides Rapid Application Development (RAD), which helps developers to shorten the development time while keeping the applications scalable and reliable. In this study, the methodology for developing ASP.NET applications using C# is introduced for developers to follow as a general guideline.

Internet Banking is one of the Web applications which is a good example of Internet computing environment and can be implemented with many programming languages such as C#, Java, PHP, ASP, and Perl. By using C# with ASP.NET, it provides the flexibility to code with an amount of rapid development. The entire processes of developing Internet Banking system by using ASP.NET are described throughout this study. This includes object-oriented approach to be applied in analysis and design phases.

IM-02-06**DESIGN AND IMPLEMENTATION OF WEB SERVICES WITH UML**

Rupesh Shrestha

Dr. Dencho N. Batanov

Web Services will transform the web from a collection of information into a distributed device of computation. In order to employ their full potential, appropriate description means for web services need to be developed. Although there are number of methods, techniques and tools for design of Web Service Definition Language available, developers still lack of methodology a systematic process to apply and use. The object modeling of web services and mapping of Unified modeling language (UML), Extensible Markup Language (XML) documents, Java classes and Web Service Definition Language (WSDL) are the key tools of this study.

In this study, the methodology provides the appropriate conceptual model for developing and describing web services and their composition. The model transformation lies at the center of the approached study. This work in progress first started with studying the WSDL specification and schema while building the first prototype model In Unified modeling language (UML). A prototype tool then applied to generate XML Schemas from UML, and the XML document is then mapped to java classes. A prototype tool is shown to generate java classes from XML documents required for deployment of Web Services.

Later in the implementation phase it is shown what the Java Application program interface (API) for XML do and how they make writing web applications easier. It also mentions various other technologies currently available and how they can be used in conjunction with the Java API for XML.

IM-02-07**PARTIALLY RECURRENT NEURAL NETWORKS FOR FILTERING AND FORECASTING**

Nguyen Thi Hoang Yen

Prof. Huynh Ngoc Phien

The applicability of partially recurrent neural networks (PRNNs) for filtering and forecasting has been tested in this study. The natural data series, namely discharge time series, in the Red River Basin in Viet Nam have been filtered and forecasted with the two selected PRNNs: the recurrent neural network (RNN) with only feedback from hidden nodes (called PRNN-I) and the Elman RNN (called PRNN-II). The on-line learning algorithm proposed by Chairatanatrai, in which the Error Back Propagation (EBP), Error Self-Recurrent (ESR), and Recursive Least Squares (RLS) methods are employed has been modified to apply to the two selected PRNNs. The obtained results indicate that the modified algorithms can be used to efficiently train the PRNN-I and PRNN-II for both forecasting and filtering purposes.

The network structure much affects the performance of network model. For the best network performance, an appropriate network structure is required. Since the issue is still under research, the trial-and-error method has been used in many applications so far. In this study, a procedure for defining the appropriate network structure based on the Bayesian Information

Criterion (BIC) has been proposed instead. The proposed procedure has been successfully applied to determine the appropriate structure for the RNNs considered in the present study.

In order to be able to compare the performance of the selected PRNNs with the fully recurrent neural network (FRNN), forecasting and filtering the flow time series in the Red River Basin with the FRNN have been carried out as well. The experimental results show that the PRNN-II model performs best in forecasting while the best performance in filtering goes with the FRNN model. The performance of the three models, the FRNN, the PRNN-I and the PRNN-II is, in fact, quite comparable in forecasting with the same mean efficiency index (EI) of 0.93 in the validation phase and only few percents difference in the training phase. With high value of EI of over 0.90 for most experimental cases, one may conclude that the three models can well forecast the stream flow in the Red River Basin. Good results of filtering the stream flow data in the Basin have also been obtained. High value of EI of over 0.90 has been obtained for most cases of using FRNN and PRNN-I for filtering. However, lower values of EI with its mean of 0.86 are found when the PRNN-II model is applied.

IM-02-08

OBJECT-ORIENTED METHODOLOGY FOR DEVELOPING WEB SITES WITH SEARCH CAPABILITIES: A CASE STUDY OF UNEP RCC.AP

Amalia S. Canullas

Dr. Dencho N. Batanov

The Internet is a vast place comprised of millions of computers sending information back and forth in packets. Many have likened the Internet to a huge, global library. While true in some aspects, it has some unique differences. There is no central "card catalog"; the Internet's growth is outpacing the ability of humans or technology to keep up with it; its sheer size is unknown and perhaps unknowable; and content is (to say the least) of uneven quality.

One of the essential challenges of the Internet is to make its value available to the millions of new users who have had no formal training or experience in query formulation or search strategies.

There are many site search tools available, for almost every platform, web server and site you can imagine but the general procedure in applying it is still a challenge to most web developers. Adding a search engine does increase the complexity of a website's infrastructure. Therefore, the promising advantages of object-oriented modeling over the traditional approach will be used to meet this challenge. With this challenge in mind, the author had decided to focus the study into developing a methodology that would help website developers in adding a site search capabilities website. The author had proposed an object-oriented model in analyzing and designing the site search engine application. Finally, this model was implemented in the case of UNEP Regional Resource Center for Asia and the Pacific (RRC.AP) website.

IM-02-09

DEVELOPMENT AND DEPLOYMENT OF WEB-BASED TRAINING SYSTEM IN AN ORGANIZATION CASE STUDY: TELEPHONE ORGANIZATION OF THAILAND

Jantana Techasirinugool

Dr. Dencho N. Batanov

The concept of globalization leads the organization compete in the world competition. The organization needs more efficiency of the system solution, which it able to increase the opportunity of the employee for self training or checking their ability to suit for the job requirement and self assessment. As the result the organization can promote the staff to the job properly and more effectively.

The study begins with reviewing and analyzing how knowledge can be shared and used to support Web-Based Training in organization using the principles of Total Knowledge Management and Enterprise Information Portal. Then using information analysis and design methods, propose a set of general procedures and models for understanding the system, which could be easily adapted to actual implementation. The model proposed in this study has already been reviewed and tested with a prototype system to demonstrate its applicability. Finally, the demonstrative application is implemented for illustrating some necessary components of the methodology as developed in this study.

IM-02-10

AUTOMATED MANAGEMENT OF STORAGE NETWORKS IN HETEROGENEOUS ENVIRONMENTS FOR INFORMATION BASE SYSTEM

Chaiwat Charuraksa

Dr. Dencho N. Batanov

Currently, the need of storage capacity is growing very fast, particularly in the enterprise computing system that it also requires continuous process for serving on a 24x7 basis. Therefore, the feature's keys of information base system are not only serving data but also providing high available and reliable infrastructure to the system. Regarding these requirements, storage network architectures have been proposed for fulfilling the demands of information base system by distributing its resources. Since the user's choice is the user's right, a system usually consists of many devices from various vendors and is implemented by several architectures. Then, many systems have been shifted into heterogeneous environments. Consequently, the management for information base system must be conducted to more complex network architectures.

Automated management of storage network in heterogeneous environments for information base system is designed by three main objectives that are simplifying storage management, reducing the cost of ownership, and improving performance of data delivery, availability, and reliability. To achieve these objectives, the following services are proposed and implemented in the

prototype system that are automated device configuration, intelligent source of information discovery, automated backup operation on schedule, and automated recovery operation from disaster.

As this system is designed for managing in heterogeneous environments that it is compound of many kinds of devices from various vendors and each device also requires for individual management, then the system's management design is divided into two components. One is a management facade component designed for manipulating the shared resource such as configuring each device. Another is a management policy component designed for defining rules to achieve the management goals that may make a change to shared resources through their management facades e.g., operating automated backup on scheduling and automated recovery from disaster.

IM-02-11

ERP AND DATA WAREHOUSE FOR DECISION SUPPORT

Badruddozza Mia

Dr. Vatcharaporn Esichaikul

This thesis contains illustrative descriptions of inconveniences of analyzing data in ERP system to provide decision-support and planning information with a comparison of doing this in data warehouse. It describes the circumstances where data warehouse works better than ERP for analyzing data. A case study has been made to see the data arrangement in ERP and data warehouse from the decision-support point of view. An enterprise information system model containing ERP and data warehouse has been proposed in order to optimize the utilization of operational data for strategic use for large and information driven organizations. Problems of transforming data from ERP along with other internal and external sources to data warehouse have been identified. Data transformation framework architecture for resolving the problems of transforming data from heterogeneous sources has been proposed in this thesis. Prototype implementation of data transformation tool has been developed on client/server architecture using SQL Server and Visual FoxPro in a thin client approach of development. Resolving mechanisms of different types of conflicts that occur during data transformation are described in this thesis.

IM-02-12

A MOVIE RECOMMENDATION SYSTEM VIA MOBILE PHONE

Wongsuda Supaporn

Dr. Vatcharaporn Esichaikul

The number of mobile phone users has quickly exceeded the number of PC users. A new market has opened up in Thailand in recent years and is both exciting and innovative, since it can reach a massive number of potential customers in real time. However, this market is still under development. For it to become more interesting, a better mobile network infrastructure is needed, along with a new Mobile Commerce portal and ways to advertise which will appeal to customers. For Mobile Commerce market, the only suitable form of advertisement can only be generated through the mobile phone itself. Hence, the key tool of this study is in supplying an effective mobile advertisement that invites customers to use WAP Services.

In this study, the methodology provides the appropriate conceptual model for developing and describing mobile advertisements and their composition. The model transformation lies at the center of the approached study. This work first started with studying the flow of information of the mobile advertisement over different networks. Then a model is constructed to show the structure of the study framework. From there, each algorithm associated with the personalization of mobile advertisements is carefully illustrated. Eventually, a schema is drawn up and a prototype is built under the concept of sending the information from the Web to the WAP model. The prototype tool is shown to generate the results of a mobile advertisement generated by Java classes from the Web server and sends the information to the Short Message Service Server and finally to the user's mobile phone. Later in the implementation phase, it shows that the user has an option of accessing the WAP Service over the short message received. Thus, the mobile advertisement promotes movies in the theater, the WAP Service and the technology.

The result of this application is a personalized mobile advertisement that satisfies both the customer and the theater promoting it. It also appeals to the mobile network service provider since the service is done over their infrastructure and half of the revenue assumes to be yielded to them. This study mentions various technologies currently available and how they can be used in conjunction with the Web and WAP services.

IM-02-13

VERTICAL JAVA FRAMEWORK FOR DEVELOPING WEB BASED TEACHING AND LEARNING SOFTWARE

Latifov Murodillo

Dr. Dencho N. Batanov

Reuse of software is considered an important issue in software engineering discipline. Reuse exists in many forms, cutting and pasting piece of code, components, design reuse, subclassing, delegation, and frameworks. Frameworks are the way that object-oriented software achieves most reuse. In this study, the methodology to develop vertical application frameworks applying object-oriented approach has been expressed as general guidance for developers to follow.

Vertical application frameworks are built to encapsulate common characteristics and behavior of objects in a domain of interest from object-oriented point of view. Educational system, particularly Question and Answer model as interaction mechanism in discussions selected as domain of underlying framework.

Methodology proposed in this work covers overall life cycle of object-oriented software development. Prototype framework is tested by applying to sample application for correctness and validity.

IM-02-14

DEVELOPMENT OF WEB-BASED EXPERT SYSTEMS CASE STUDY WEB-BASED ADVISER FOR INSURANCE PRODUCT SELECTION

Benjamard Nittayapat

Dr. Dencho N. Batanov

This thesis study deals the analysis, design and implementation of methodology in development of Web-based expert systems. Conduct case study Web-based adviser for insurance product selection. Nowadays, in e-commerce application, sometime potential buyers may be interested in receiving recommendations about what to purchase. The mainstream of Web-based expert system is rules engine. Recently server-based expert system approach is proposed.

In this paper, I present a Web-based expert system adviser. The knowledge base of product domain, business rules and inference engine is implemented by JESS and Java servlet. The system gathers the user's requirements on a particular insurance product by questioning the user and consulting expert to find the insurance product that best meet the user's requirement with recommender underwriting process. For the expert system, the design of JAVA classes and its functions are discussed. As a sample implementation, insurance product selection Web-based expert adviser is built where the system tries to find an insurance plan that meet user's requirement best.

IM-02-15

A HYBRID APPROACH TO GENERATING CROSS-SELL RECOMMENDATIONS

Pham Thi Phuong

Dr. Peter Haddawy

The development of Internet and e-commerce gives consumers a lot of options to select among products or services available in virtual stores. However, consumers have to struggle with the glut of product and service information to make purchasing decisions. For this reason, giving personalized and cross-selling recommendations becomes important for e-commerce sites. They help online stores increase sales while eliminating information overload and gaining better customer satisfaction. Most shopping sites give recommendations for products or services within a category such as book, or movie. This thesis presents an algorithm to generate cross-selling recommendations where products and services that might belong to different categories by combining two algorithms: K-means clustering and *a priori* association rules.

The experiments with simulated data on the prototype show that generated recommendations are reasonable with an appropriate set of parameters. The cross-category characteristic makes the presented algorithm a nice compliment to existing single-category recommender systems. The two types of algorithm can be used together in shopping sites that sell many different types of products and services.

IM-02-16

AN E-CATALOG MANAGEMENT FRAMEWORK FOR BARTER TRADE EXCHANGES

Le Anh Vu

Dr. Peter Haddawy

Barter is developing very fast recently and it has grown as a multibillion-dollar industry. Barter exchanges can help businesses solve many problems such as short of cash, slow inventory by allowing them to exchange products and service with each other without using money. As intermediaries, barter exchanges need an effective way to facilitate information exchange between members and encourage trading activities. This thesis provides a framework that supports barter exchange operations. The operations are considered as a process of creating, integrating, and disseminating electronic catalogs. The framework uses XML and classification standards such as ebXML, UNSPSC to ease information exchange process. In addition, it uses web services for system integration. The evaluation of the prototype shows that the framework successfully satisfies the above requirements; it also gain very positive responses from the barter professionals.