Depatetment of Computer Engineering & IT
ABSTRACTS OF PROJECTS FROM B. TECH. (Information Technology)

RISK ESTIMATION USING NEGATIVE BINOMIAL REGRESSION

Radiation has a great impact on human body. Excess exposure to ionizing radiation can have adverse effects. The effects of radiation may even be carried out in the next generation in some instances. Even at low levels it can destroy and break down important cells.So, it is very necessary to study the effects of exposure to radiation on human body. Various risk models exists that have done study in this field. However, most of them have have done risk estimation for cancer or leukemia. Very less risk estimation mod- els have been found that can state the risk of such radiation exposure at various levels that causes cardiovascular diseases. So, we have made a risk estimation model using Nega- tive Binomial Regression that can estimate the relative risk of moratlity of exposed person.
Two models, namely multiplcative and linear models, have been built for stroke and heart diseases using the data of people, who were exposed to Atomic Bomb explosion in Hiroshima and Nagasaki, and various risks estimations, for this dataset, have been made. As per our model, we have found that doses above 1.25 Gy are associated with an elevated risk of cardiovascular diseases, but the degree of risk at lower doses is unclear.
Radiation has a great impact on human body. Excess exposure to ionizing radiation can have adverse effects. The effects of radiation may even be carried out in the next generation in some instances. Even at low levels it can destroy and break down important cells.So, it is very necessary to study the effects of exposure to radiation on human body. Various risk models exists that have done study in this field. However, most of them have have done risk estimation for cancer or leukemia. Very less risk estimation mod-els have been found that can state the risk of such radiation exposure at various levels that causes cardiovascular diseases. So, we have made a risk estimation model using Negative Binomial Regression that can estimate the relative risk of moratlity of exposed person. Two models, namely multiplcative and linear models, have been built for stroke and heart diseases using the data of people, who were exposed to Atomic Bomb explosion in Hiroshima and Nagasaki, and various risks estimations, for this dataset, have been made. As per our model, we have found that doses above 1.25 Gy are associated with an elevated risk of cardiovascular diseases, but the degree of risk at lower doses is unclear.Radiation has a great impact on human body. Excess exposure to ionizing radiation can have adverse effects. The effects of radiation may even be carried out in the next generation in some instances. Even at low levels it can destroy and break down important cells.So, it is very necessary to study the effects of exposure to radiation on human body. Various risk models exists that have done study in this field. However, most of them have have done risk estimation for cancer or leukemia. Very less risk estimation mod-
els have been found that can state the risk of such radiation exposure at various levels that causes cardiovascular diseases. So, we have made a risk estimation model using Negative Binomial Regression that can estimate the relative risk of moratlity of exposed person. Two models, namely multiplcative and linear models, have been built for stroke and heart diseases using the data of people, who were exposed to Atomic Bomb explosion in Hiroshima and Nagasaki, and various risks estimations, for this dataset, have been made. As per our model, we have found that doses above 1.25 Gy are associated with an elevated risk of cardiovascular diseases, but the degree of risk at lower doses is unclear.

Supporting Efficient and Scalable Multicasting over Mobile Ad Hoc Networks

Group communications are important in Mobile Ad hoc Networks (MANET). Multicast is an efficient method for implementing group communications. However, it is challenging to implement efficient and scalable multicast in MANET due to the difficulty in group membership management and multicast packet forwarding over a dynamic topology.We propose a novel Efficient Geographic Multicast Protocol (EGMP). EGMP uses a virtual-zone-based structure to implement scalable and efficient group membership management. A network-wide zone-based bi-directional tree is constructed to achieve more efficient membership management and multicast delivery. The position information is used to guide the zone structure building, multicast tree construction and multicast packet for- warding, which efficiently reduces the overhead for route searching and tree structure maintenance. Several strategies have been proposed to further improve the efficiency of the protocol, for example, introducing the concept of zone depth for building an optimal tree structure and integrating the location search of group members with the hierar-chical group membership management. Finally, we design a scheme to handle emptyzone problem faced by most routing protocols using a zone structure. The scalability and the efficiency of EGMP are evaluated through simulations and quantitative analysis. Our simulation results demonstrate that EGMP has high packet delivery ratio, and low control overhead and multicast group joining delay under all test scenarios, and is scalable to both group size and network size. Compared to Scalable Position-Based Mul-ticast (SPBM) [20], EGMP has significantly lower control overhead, data transmission overhead, and multicast group joining delay.

Decentralized Social Networking Using the Open Source Project Diaspora*

Online Social Networks boast of several millions of users. To enhance user interaction and provide better user experience, many applications are built on them that encourage different forms of social network sharing. The heart of such social networking applications is user identity and profile, their social graphs and connections, interactions and generated content. Centralized Social Networks behave as 'information silos', where all this huge user generated data is stored in massive central databases. These Centralized Service Providers, acting as sole control entities having uncontrolled access to over millions of user data, often disclose them to third parties and have profitable data mining activities running on them to exploit it for their own commercial purposes.

AN EFFICIENT INTER-VM COMMUNICATION MODEL SUPPORTING LIVE MIGRATION

Virtualization technology in cloud environment is characterized by the property of sharing hardware resources among multiple operating systems. It also maintains isolation between the virtual machines. In virtualized environment, network intensive applications such as web services or databases are being consolidated onto a single host. In these applications, an efficient communication between virtual machines(VMs) is required. Isolation is very important for security reasons but, it increases inter-VM communication overheads. The communication between VMs takes place through shared memory or network datapath. In this project, we are extending the prior work done in shared memory based commu-nication in KVM. We have proposed a protocol which bypasses network stack completely and establishes communication using shared memory when the VMs are co-located. Due to load balancing policies, the VMs in communication may be migrated to same or dif-ferent hosts. Hence, the proposed protocol also identifies co-located and non-co-located VMs after migration. Our protocol provides transparency with respect to live migration. The

protocol also provides synchronization between communicating VMs and achieveing higher throughout, improves inter-VM communication performance.

Android Application Development Sophistication and Assistance tool

ADT(Android development tool) is the only platform where user can develop his mobile application from scratch. Focusing on this issue, we made an objective to provide user most of source codes of existing android mobile applications in order to assist him. For achieving this, we use the concept of plug-ins ,that can be add to the ADT ,and also can update .One of this plug-ins will provide user the database containing .apk, which contains all files related to application. Also with other plug-in ,user can upload his own code ,that will be available to the other users publicly and also can review. We also provide other plug-ins giving information about android blog connect,news and updates. With this additional features, developer can use existing code, which can provide guide or hint to him while developing his own app. Also his code can be published and entered to the database ,so that his code will be accessible to the other users.

PERSONAL ASSISTANT APPLICATION USING ANDROID

Aim of this project is to create an android application that will help user to manage his meetings efficiently on the android mobile phones. While most of the inbuilt notes and to-do applications provide user to manage their meetings, our application handles meeting topics, their venues, and automatically deals with the switching of mobile profile from general to silent. Following are the main features of our application.

1. Management of Meetings:
User can save the meetings along with the topic, venue, date and time, the persons that are to be invited to the meeting.
2. Handling of meeting cancellation:
When the person arranging the meeting wishes to cancel it, then he will just have to enter the meeting-id and press the CANCEL button. The corresponding message of cancellation will be sent to concerned people.
3. Profile Switch:
When the user will enter into that specific location saved in his meeting or will be around in the radius of 50 meter, he will get the alert that he is near the venue of scheduled meeting and the profile will be automatically switched from GENERAL to SILENT. This is achieved using GPS.
4. Automtically adding the received message as a arranged meeting in receivers mobile phone so that he will also get the required notifications about it.
5. Sending message:
User can send text message to any of his contacts, using the same application.
6. Currency Converter:
This application provides currency conversion, based on the current conversion val-ues. This seems useful for business related people. Though there are some android applications that handle this meetings related task, they are not as efficient as our application. For instance, our application not only main- tains the meetings, but it also enables user to take the appropriate action when meeting get cancelled, it automatically sends message to the concerned people. Also, it provides automatic profile switch and provides respective alert on mobile whenever user enters into the meeting location or is around the location, considering the periphery as of 50 meter. The development environment consists of Windows 7 OS, Eclipse Java IDE with Android plug-ins, Sun Java SE Development Kit, Android 4.0.1 operating system, Android framework API level 14 and Android based handset.

LOCATION BASED SEARCH WITH REAL TIME TRAFFIC SYSTEM

Mobile phones never have been more popular, and powerful smart phones are now have become regular choice for customers. Stylish and versatile phones have hardware features like GPS, accelerometers, and touch screens that have enticing platform upon which we can create innovative mobile applications. Mobile computing systems are playing a more and more important role in everyday life. Android has opened mobile phone development to thousands of developers who havent had access to tools for building mobile application. Experienced mobile devel-opers can now expand into Android platform, leveraging the unique features to enhance existing products or create innovative new ones. Applications based on both local and wide distributed wireless networks are being launched all the time. It is always observed that when travelling to a new city or even in known cities, we face many delays. These delays are due to lack of information of roads in new city, lack of information about road condition and some road blocks due to road works, accidents, festivals or rallies and lack of information about current traffic condition of the roads. It is also observed that many delays can be minimizes if we know all the factors discussed above. Many location based search apps are in the markets but many of them dont use real time data in order to calculate best suitable path to destination. These include the current accidents details, traffic delay reasons so that user can easily decide best suitable path to destination. We have used Google Places Api, Google Maps Api, Google Direction Api for providing location based search depending on users choice and for real time traffic route suggestion we have used Real Time Vehicle detection algorithm in order to calculate the number of vehicles based on the road surveillance video and using this, congestion on road can be computed and with the help of Linear regression equation and Dijkstras algorithm best suitable path can be calculated.

MULTIPURPOSE ANDROID APPLICATION FOR VISUALLY CHALLENGED USERS USING BRAILLE ALPHABETS

Mobile phones have become a necessity in people's life today as almost everyone will have at least one with them. Short message service (SMS) which is an important component in mobile phones is growing popular among mobile phone subscribers. The use of SMS nowadays is more than the calling function itself. In order to make the SMS function in mobile phones more efficient and attractive for visually challenged users, the Braille language is used as a mode of input. This Braille language is used by blind people for reading their books. Same input format is used for this application to make in convenient for blind users. This Application consists for three main modules. First one is used to send SMS to any user, in which user enters the message in the form of Braille language. Second module is used to receive the messages from any sender and readout that message to the recipient. Third module is used to track current location of user using GPS receiver in their phones. Later on it can be used to calculate the distance between current location and destination or any two locations, in which you can provide the input using Braille language.

CONTROLLING IP SPOOFING USING INTER-DOMAIN PACKET FILTERS USING BGP UPDATES

In TCP/IP protocol IP header contains source and destination addresses. Ensuring that an IP packet carries a correct source address would be valuable for many purposes.Services that rely on correct source addresses (congestion control, fair queuing, and source based traffic control schemes) would profit. Network problem diagnosis are now able to locate the possible sources of a problem and could also besimplified. Moreover, this would assist in solving one of the most important problems

in network security: attackers commonly forge source addresses to avoid responsibility for their malicious packets. Examples include DDoS attacks, TCP SYN flooding attacks, and smurf attacks. Reliably locating the attacker has been difficult because defenders cannot easily verify that a packet was indeed sent by the node specified in its source address. The IP Spoofing is a critical threat to the valid use of the Internet throughout the world. All impediment mechanisms employed have been prevented by the capacity of attackers to invent or spoof the source addresses of the IP packets. By using IP spoofing, hackers can avoid discovery and for policing the attack packets put more burden on the destination network. Here, we suggest an Inter-Domain Packet Filter (IDPF) architecture which can minimize the level of IP spoofing on the Internet. The important point of this is we dont need any Global Routing information .The IDPFs are implemented on the network border routers and build from the information gathered from Border Gateway Protocol (BGP) route updates. Here the IDPF framework works in such a way that it will not block packets with legitimate source addresses. At the end of this thesis we show that even with fragmentary implementation of the IDPFs on the Internet we can reduce the spoofing capacity of attackers and they can help in localize the source of an attack packet to a fewer number of candidate networks.

3D PASSWORD FOR ATM ALONG WITH FACE RECOGINITION

In our project, we propose and evaluate our contribution in a new scheme of authen- tication in ATM. This scheme is based on a face recognition followed by virtual 3D environment. In this project, PCA based face recognition is used. 3D passwords which are more customizable and interesting way of authentication. The 3D password is a multi factor authentication scheme. Users navigate through the Virtual environment. The com-bination of all interactions towards the items and towards the virtual 3D environment constructs the users 3D password.

STATIC PROGRAM ANALYSIS OF C

Static analysis examines code in the absence of input data and with-out running the code, and can detect potential security violations run-time errors (e.g. dereferencing a null pointer) and logical inconsistencies. Computer program analysis is the process of automatically analysing the behaviour of computer programs. There are two main approaches in program analysis:
1.Static Program Analysis.
2.Dynamic Program Analysis.
We are concerned about the static analysis of c program. Static code analysis is the process of detecting errors and defects in software's source code. Static analysis can be viewed as an automated code review process. Code review is one of the oldest and safest methods of defect detection. It deals with joint attentive reading of the source code and giving recommendations on how to improve it. This process reveals errors or code fragments that can become errors in future. It is also considered that the code's author should not give explanations on how a certain program part works.
The program's execution algorithm should be clear directly from the program text and comments. If it is not so, the code needs improving. The code review usually works well because programmers can notice errors in somebody else's code much easier than in their owns. The only crucial disadvantage of the joint code review method is an extremely high price: you need to gather several programmers at regular times to review a fresh code or re-review a code after recommended changes have been applied to it. The programmers also need to have a rest regularly, as their attention might quickly weaken if they review large code fragments at a time, so there will be no use of code review then. Static code analysis tools are a compromise solution. They can tirelessly handle source textsof programs and give recommendations to the programmer on what code

fragments he/she should consider. Of course, a program can never replace complete code review performed by a team of programmers, but the ratio use/price makes usage of static analysis a rather good practice exploited by many companies. Static program analysis aims at determining properties of the behaviour of a program without actually executing it. It is the process of detecting errors and defects in software's source code. Static analysis can be viewed as an automated code review process.

GEOGRAPHICAL INFORMATION SYSTEM

In the past, Geographical Information has been stored and presented in the form of maps. Today There is trend of using computer based GIS which shows maps digitally. Although many GIS have been successfully implemented, this is our best approach to implement GIS application by extending its capabilities of presenting geographic and other information. Geographical Information System i.e. GIS is used today worldwide. It is used to find specific location on map i.e. used to find lattitude and longitude of perticular place that user want to find. Our Project intend to show how GIS works and how it is useful in different areas. This paper also identifies some of the key areas where this GIS systems could be very useful. In our project, we are using spatial data and displaying information graphically.

BLUETOOTH STETHOSCOPE FOR TELEDIAGNOSIS AND VIRTUAL CLINIC

Bluetooth is a technology that enables a device to transfer data wirelessly over a small network. This feature, combined with the technology of internetworking, it is possible to create a virtual environment wherein the doctor remotely diagnose a patient living in a remote location and having internet connectivity. He can accurately measure his patients heartbeats, listen to them, diagnose his condition and prescribe the medicine needed, all this, without having to move from his clinic. The project requires a Bluetooth enabled stethoscope along with the patient. The stethoscope record the patients heartbeats with the help of its sensing element (piezo sensor), perform digital signal processing (DSP) on it, and send the data to the bluetooth receiver connected to a computer. The computer can then produce audio signal which can be transmitted to another computer located far away, with the help of real time transfer protocol (RTP). The Destination computer can then simulate the heartbeats for the doctor’s diagnosis. In addition to this, a database concerning the patient details and his history and medicinal diagnosis will be available to the doctor for any reference needed. The project can therefore, help create a new environment of patient diagnosis in a virtual surrounding, along with a database always available, if the need arises. The doctor can keep a track of his patient and the patient can remain in contact with him, maintaining a harmonious and detailed relation.

Student Academic Performance Indicator

Data mining is the process of digging through huge amount of data to discover rules and patterns from data. Some types of knowledge discovered from a database can be represented by a set of rules. Other types of knowledge are represented by equations relating different variables to each other or by other mechanisms for predicting outcomes when the value of variables is known. In our project we have predicted the 12th standard board exam marks of a student from his full portion class tests marks and the number of hours he has studied for each subject. Using these predicted marks and few psychology questions we will guide him to choose the field that best suits him.

CONTEXT-BASED ANOMALY INTRUSION DETECTION SYSTEM

Anomaly detection approaches are generally efficient in detecting new attacks. How-ever, they fail in providing any further information regarding the nature of attacks. The first contribution of this paper is to equip an anomaly detection approach with a diagnosis module that classifies anomaly approach outputs in one among well-known attack cate-gories. The second contribution concerns a context-based definition of normal network traffic profiles. Experimental studies shows, considering normal profile for each service provides better results than considering a unique global normal profile.

CO-OPERATIVE WIRELESS INTRUSION DETECTION SYSTEM USING SNMP MIBs

In emerging technology of Internet, security issues are becoming more challenging. In case of wired LAN it is somewhat in control, but in case of wireless networks due to exponential growth in attacks, it has made difficult to detect such security loopholes. Wireless network security is being addressed using firewalls, encryption techniques and wired IDS methods. But the approaches which were used in wired network were not successful in producing effective results for wireless networks. Many attempts were made to secure wireless ad hoc networks, but due to their special ad hoc nature and strict constraints, finding an optimal and comprehensive security solution is still a research challenge.It is so because of features of wireless network such as open medium, dynamic changing topology, co-operative algorithms, lack of centralized monitoring and management point, and lack of a clear line of defense. So, there is need for new approach which will efficiently detect intrusion in wireless network. Efficiency can be achieved by implementing co-operative IDS. Here, we have proposed a co-operative anomaly and network-based IDS using SNMP MIBs for wireless ad hoc networks. Our proposed system makes use of threshold scheme and detects the DoS attack. Using MIB datasets collected from real experiments involving a DoS attack, we validate the possibility of our approach.

VeracityNXT DX11 Media Validation Framework Development

Microsoft DirectX is a collection of application programming interfaces (APIs) for han-dling tasks related to multimedia, especially game programming and video, on Microsoft platforms. Newly launched Microsoft OS, Windows 8, comes with DirectX11.1 API, which provides new framework for video playback with enhanced capabilities and new technologies. All applications on Win8,using hardware acceleration, need to interact with hardware, and they do this through a device driver. Hardware manufacturers have to write these drivers for a particular DirectX version's device driver interface (DDI), and test each individual piece of hardware to make them DirectX compatible. DX11.1 video path being newly introduced and driver code being such a huge code, any change or extension to the driver code is susceptible to regressions or compatibility issues. Hence, any extension or addition to the code needs to be tested thoroughly before actually deploying it as a device driver. This report describes the details of development of such automated framework, VeracityNXT, which will provide a simplified platform for development of Dx11 tests.The framework aims at contributing to Nvidia's Driver Validation System (DVS) in order to ensure correctness of DX11 pipeline on Nvidiaplatforms .

A Self-configurable New Generation Children Tracking System based on Android Mobile

Terminals

It is software that allows parents to monitor their child's cell phone. All incoming and outgoing calls, texts messages can be seen by the parents, who can also monitor where their children are (through GPS), access a history of where they've been and set up alerts if their children are going outside of approved geographical zones. The information like the details of incoming call, text and multimedia messages and the timely location update of their children are also stored in the centralized server. Parents may later login into the centralized server and view the details of their child‘s mobile usage.

HYBRID INTRUSION DETECTION

Intrusions detection systems (IDSs) are systems that try to detect attacks as theyoccur or fter the attacks took place. IDSs collect network traffic information from some point on the network or computer system and then use this information to secure the network. Intrusion detection systems can be misuse-detectin or anomaly detection based. Misuse-detection based IDSs can only detect known attacks whereas anomaly detection based IDSs can also detect new attacks by using heuristic methods. In this paper we propose a hybrid IDS by combining the two approaches in one system. The hybrid IDS is obtained by combining packet header anomaly detection (PHAD) which is anomally-based IDS with the misuse-based IDS Snort which is an open-source project. The hybrid IDS obtained is evaluated using the MIT Lincoln Laboratories network traffic data (IDEVAL) as a testbed. Evaluation compares the number of attacks detected by misusebased IDS on its own, with the hybrid IDS obtained combining anomaly-based and misusebased IDSs and shows that the hybrid IDS is a more powerful system.

CAMPUS CONNECT

The Mobile platform has surpassed the Desktop platform to become the prime mode of communication,entertainment and business.Today, there are 750 million active Android devices across the globe. The samrtphone market is rapidly expanding and innovat-ing. With the advent of new Operating Systems and technologies everyday, the focus of developers has shifted from desktop applications to mobile applications.Being able to perform complex computations,day-to-day activities alongwith entertainment on your smartphones has fuelled the smartphone market. We proposed an application to ease the life of students by providing a better platform for communication and scheduling.Our project involves the development of one such Third party Android Application, which seamlessly connects all student activities on a college campus on a mobile platform. These activities involve over the air synchronization of class and exam timetable,as well as a discussion forum for student interaction at a single, portable platform.

Sensemaker: Automatic Word Sense Discovery

Manually compiled dictionaries have some disadvantages like, they contain rare senses but miss out corpus/domain specific senses (other disadvantages are discussed later). Word Sense Induction (WSI) is the task of identifying the different senses (uses) of a target word in a given text.[2] Many clustering and graph based approaches have been applied to solve the problem of WSI. The existing solutions are very complex. The algorithms which we have designed(Minimum Spanning Tree based Clustering and Average Linkage Clustering), have a simpler ap-proach in which only the context words of the target word are considered in order to obtain clusters representing different senses of the target word. We have carried out various experiments by varying the values of the

parameters which influence the quality of the clusters obtained. We have observed that the results are promising. We have compared both the algorithms based on their performance.

Integrating Multi-label Classification And Novel Class Detection For Concept Drifting Data Streams

In the conventional method of single-label classification every example is associated with a single class label and a classifier learns to associate every new test example with one of these known class labels. When each example maybe associated with multiple labels, then the method is known as multi-label classification. It is a standard assumption that the total number of classes present in a typical data stream classification task is fixed. This assumption maybe invalid in a real streaming environment, in those cases where new classes may occur in the stream. Conventional data stream classification techniques cannot recognize novel class in-stances until it is done manually, and labeled instances of that class are used for training the learning algorithm. The presence of concept drift makes this problem even more challenging, when the underlying data distribution changes over time and if the data is multi-labeled where multiple target labels must be assigned to each instance. We pro-pose a novel technique that can detect the emergence of a novel class in the presence of concept-drift for multi-labeled data.

Risk Estimation Models for Atomic Bomb Survivors

Objective:Applying suitable methods to gain insight in the model developed by RERF and to develop the best possible risk estimate for exposure to low-dose, low linear energy transfer (LET) radiation in human subjects. The study of the distribution and determinants of disease due to ionizing radiation prevalence is done primarily in Radiation Epidemiology. Epidemiologists seek to relate risk of disease to different levels and patterns of radiation exposure. In our project we examine statistical model of Poisson regression previously employed for the estimation of radiation risk and propose model of Hurdles which could improve the estimation of risk. We examine different regression techniques, which overcome the underlying assumptions of Poisson Regression for risk estimation and propose Hurdle's Model for the same. The models need application of logarithmic transform to yield the additive model instead of multiplicative model, which is usually used in Risk Assessment and thus obtain the Linear Dose-Response Model.

Handling Uneven Length Text Documents and Modeling Poisson Regression in Multinomial Naive Bayes

While Naive Bayes is quite efficient in various text mining tasks, it at times fails in the automatic text classification problem. Based on the observation of Naive Bayes for the natural language text, we found a serious problem in the parameter estimation process and performance due to uneven document size. We are proposing a algorithm to handle uneven document size and a new model for text classification termed as Poisson Naive Bayes Classifier which overcomes the problem in the parameter estimation process. Our new algorithm overcomes the worst effect of uneven document size and provides quite effective result. Our new classifier models the terms with respect to every document as well as with respect to every class. We consider mean over the corpus for the terms rather than considering only term-weight in the document. We also combine feature weighting with feature selection method which gives more accurate results than obtained using only feature selection. The experimental results obtained by using our proposed approach outcomes the performance of the Naive Bayes Multinomial.