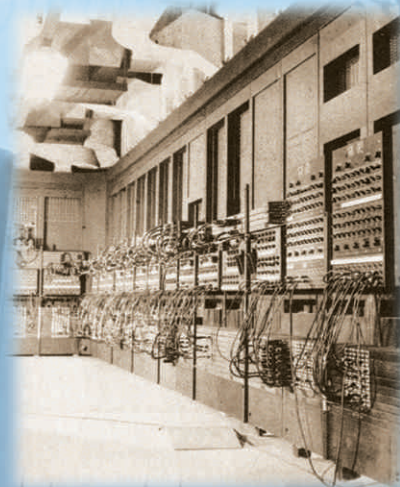
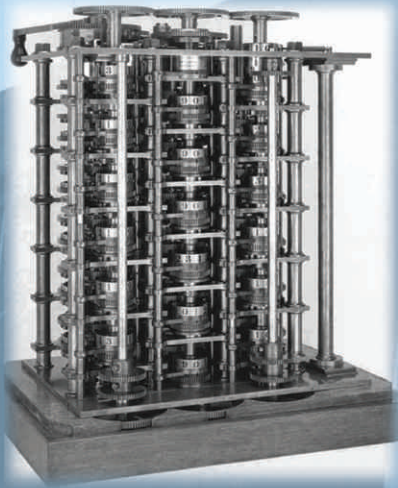


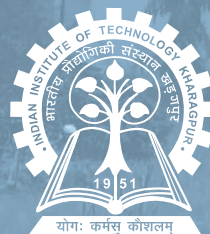
RESEARCH SCHOLARS' DAY

DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING

February 6, 2010



DEPARTMENT OF
COMPUTER SCIENCE
AND
ENGINEERING



INDIAN INSTITUTE OF TECHNOLOGY KHARAGPUR



Department of Computer Science & Engineering

The Department of Computer Science & Engineering was initiated in 1980 and the first B. Tech. batch graduated in 1982. Apart from being the department producing the first batch of graduates in Computer Science and Engineering amongst the Indian Institutes of Technology, this is one of the most reputed centres for Computer Science education and research in the country.

The hallmarks of the department are in the breadth of its academic curricula and diversity in fundamental research and industrial collaborations. Collaborative research is ongoing with researchers in internationally acclaimed universities and research institutions abroad and in India such as UCLA, USC, TIFR Mumbai, ISI Kolkata, RRI Bangalore, Perimeter Institute of Theoretical Physics, and SAC Bangalore. The Department has long-term research partnerships with leading companies such as Intel, National Semiconductors, Microsoft, General Motors, Synopsys, Sun Microsystems and Texas Instruments.

The alumni of this department are well established all over the globe achieving excellence in professional fields as well as in academics and research, and holding positions of rare distinction in leading industries and academic institutions of the world.



Message from the head

It is indeed a matter of great pleasure for me that my department is going to host the Research Scholar's Day on 6th February, 2010. On this day, the research scholars of the department will present their research activities in the form of technical presentations and posters. This will not only give them an opportunity to demonstrate their latest research findings to peers and colleagues in the institute but also act as a forum to highlight the diverse research activities of the department. Interactions with students and faculties will surely motivate the scholars to develop and improve their research work.

The success of an event like this depends largely on the collective participation and I note with great satisfaction how every faculty and research student of the department have worked hand in hand to make the event successful. I congratulate them all for their effort and wish them and the event a grand success.

Indranil Sen Gupta

List of Current Ph.D. Scholars

Arnab Sarkar	R Rajendra Prasath
Bibhas Ghosal	Rajib Maiti
Bivas Mitra	Rajiv Ranjan Suman
Bodhisatwa Majumdar	Rishiraj Saha Roy
Chandan Karfa	Sanjay Chatterji
Chester Rebeiro	Santosh Ghosh
Chhabi Rani Panigrahi	Saptarshi Ghosh
Debi Prosad Dogra	Satya Gautam Vadlamudi
Dinesh Dash	Shyamosree Pal
ESF Najumudheen	Sk. Subidh Ali
Gopal Paul	Soma Saha
Ishani Chakraborty	Soumen Bag
Joydeep Chandra	Soumyadip Banerjee
Kallol Mallick	Soumyajit Dey
Kamalesh Ghosh	Sourav Dandapat
Mahesh Raghunath Shirole	Sourav Das
Manjira Sinha	Srobona Mitra
Manoj Dixit	Subhankar Mukherjee
Maunendra Sankar Desarkar	Subhasish Dhal
Plaban Kumar Bhowmick	Subhendu Bhadra
Prasenjit Mandal	Subrata Nandi
Pravanjan Choudhury	Sudip Roy
Priyankar Ghosh	Sumanta Pyne
Rajarshi Pal	Tapas Samanta



RESEARCH ABSTRACTS



Arnab Sarkar

Email: arnab@cse.iitkgp.ernet.in

Arnab Sarkar received the B.Sc. degree in Computer Science in 2000 and B.Tech degree in Information Technology in 2003 from University of Calcutta, Kolkata, India. He received the M.S. degree in Computer Science and Engineering at the Indian Institute of Technology (IIT), Kharagpur, India in 2006 and is currently pursuing his PhD in the same institute. He received the National Doctoral Fellowship (NDF) from AICTE, Ministry of HRD, Govt. of India, in 2006 and the MSR India PhD fellowship from Microsoft Research Lab India, in 2007. He is currently pursuing his research as a Microsoft Research Fellow. His current research interests include real-time scheduling, system software for embedded systems and computer architectures.

Supervisor: Prof. Sujoy Ghose

Low Overhead Real-time Proportional Fair Scheduling

The plethora of different types of embedded systems today has initiated the emergence of various complex real-time applications which require operating under stringent performance and resource constraints. Today's power-constrained hand held devices simultaneously executing a mix of applications like real-time signal processing, Continuous media (audio and video streams), email, web browsing, etc. provide an interesting example. Another example is provided by the automotive control systems which concurrently execute a mix of hard, soft and non real-time applications on a distributed platform composed of heterogeneous multi-processors.

In recent years, proportional fair scheduling (Pfair, ERfair, Completely Fair Scheduler (CFS), etc.) is being accepted as an effective resource management strategy for the integrated scheduling of these different applications with various degrees of timeliness criticality. This is primarily because of its ability to provide temporal isolation to each client task from the ill-effects of other "misbehaving" tasks attempting to execute for more than their prescribed shares of a resource. Moreover, many applications such as multimedia audio and video streams not only demand meeting deadlines, but also demand CPU reservation to ensure a minimum guaranteed Quality of Service (QoS). These demands are of the form: reserve X units of time for application A out of every Y time units. Proportional fair schedulers with their ability to provide well-defined rate specifications form a more flexible and suitable scheduling strategy in the design of these systems.

However, in spite of its theoretical importance and usefulness, actual implementations of these fair schedulers are limited mainly due to the high scheduling, inter-processor task migration and cache-miss related overheads incurred by them. Scheduling overheads refer to the delay incurred in selecting the next task for execution. As proportional fair schedulers generally need to sort operation deadlines of tasks, they suffer scheduling complexities that are at least logarithmic in the number of tasks. Migration related overheads refer to the time spent by the operating system to transfer the complete state of a thread from the processor where it had been executing to the processor where it will execute next after a migration. Obviously, the more loosely-coupled a system, the higher will be this overhead.

Cache-miss related overheads refer to the delay suffered by resumed threads of execution due to compulsory and conflict misses while populating the caches with their evicted working sets. Therefore, processors are affined to tasks whose working sets currently exist in cache and are valid (non-dirty) since their execution results in cache hits. Obviously, the more recently a task was executed in a particular processor, higher is the probability of its working set to be present in that processor's cache. Proportional fair schedulers are usually ignorant of the affinities between tasks and their executing processors which may cause unrestricted inter-processor task migrations and heavy cache-misses.

This research endeavors to develop a framework for fast, flexible algorithms that can work effectively under a variety of practical situations like limited power, overload, faults, etc. over a wide range of architectures. The proposed fair scheduling framework is based on the following four principal mechanisms:

- (1) Frame-based Scheduling (periodic partitioning followed by global resynchronization) to guarantee bounded fairness in $O(1)$ time and restrict migrations,
- (2) Intelligent Partition-Merge Techniques to minimize global scheduling and migrations
- (3) Effective use of Processor Slack to manage power, overloads, faults, etc. and
- (4) Awareness of Task-to-Processor Mutual Affinity for overhead management.

The methods are founded on a set of theoretical bounds and experimental analysis to provide scope for developing application-specific schedulers under various fairness-speed-power-overload requirements.

We have been able to remove the $O(\lg n)$ scheduling complexity barrier of typical proportional fair schedulers by the $O(1)$ frame-based proportional fair algorithm Frame Based Proportional Round-Robin (FBPRR) which provides good bounded fairness guarantees. The multiprocessor counterpart of this algorithm called Partition Oriented Frame Based Fair Scheduler (POFBFS) is able to reduce the number of migrations in current ERfair schedulers by upto 100 times. Proportional round-robin scheduling strategies and clustering techniques have been employed within frames to further improve the fairness properties of these schedulers. We have developed Sticky-ERfair, a strictly ERfair global algorithm that reduces both the number of migrations and cache-misses by upto 40 times. Work on a partition-merge scheduling strategy with strict ERfairness guarantees is in progress. Our algorithm ERfair Scheduling with Processor Shutdown (ESPS) attempts to reduce processor energy consumption in ERfair systems by locally maximizing processor shut-down intervals through a novel procrastination scheme. Currently, we are developing its multi-processor and migration aware versions. We have also developed overload handling strategies for Pfair and ERfair multi-processor systems using efficient low-overhead admission control. Development of a fault tolerance methodology to handle single transient software faults in ERfair systems is in progress. As a future plan, we intend to develop fault-tolerant techniques (mainly for the automotive domain) to handle distributed and heterogeneous multi-processor systems by appropriate static characterization followed by partition-oriented on-line scheduling.



Bivas Mitra

Email: bivasm@cse.iitkgp.ernet.in

***Bivas Mitra** received his B.Tech. from Haldia Institute of Technology, Vidyasagar University in 2001, and M.Tech. from IIT Kharagpur in 2003 both in Computer Science and Engineering. During February 2003 to January 2006, he worked as a lecturer in the department of Computer Science and Engineering at Haldia Institute of Technology. In January 2006, he joined as a research scholar in the department of Computer Science and Engineering, IIT Kharagpur. In his PhD tenure, he has received various fellowships like national doctoral fellowship, SAP Labs India doctoral fellowship etc. and several student travel grants to participate in different international conferences. His research interests include peer-to-peer networks, complex networks, networks modeling, optical networks, wireless internet etc.*

Supervisors: Prof. Niloy Ganguly and Prof. Sujoy Ghose

Analyzing the Dynamics behind the Emergence of Stable Peer-to-Peer Networks

In this research, we primarily focus on understanding the dynamics of large scale peer-to-peer (p2p) networks. Two different aspects, namely 'network stability' and 'network emergence' are taken into consideration.

Network stability: Nodes in the peer-to-peer network join and leave the network frequently without any central coordination. This churn of peer nodes can partition the network into smaller components and can breakdown the communication among peers. The stability of the network can also get severely affected through intended attack targeted towards the important peers. Denial of Service (DoS) attack, eclipse attack, sybil attack, file poisoning are some of the important attacks that affect the stability of the p2p networks. Hence performing comprehensive theoretical analysis of stability of the peer-to-peer networks against peer churn and attack and subsequently deciding upon the optimal network is an important research problem.

Network emergence: Peer-to-peer networks are formed mainly as a result of the bootstrapping protocols followed by incoming nodes, peer churn and rewiring of the existing links. Discovering the relationship between the node dynamics and emergence of superpeer networks is an interesting and nontrivial research problem. The classical studies have shown the emergence of scale free networks as a result of the preferential attachment of incoming nodes with the 'good' existing nodes. However, we observe that the superpeer networks follow bimodal degree distribution that sharply deviates from the power law behavior of scale free networks. Hence the classical theories related to preferential attachment and its variations need to be modified to explain such deviation.

Methodologies followed to analyze network stability and to explain the emergence of superpeer networks are mainly drawn from complex network theory.

Stability analysis: Our research concentrates on the application of classical concepts of statistical mechanics like percolation theory, generating function formalism, condensation theory etc to develop an analytical framework to measure stability of p2p networks. We characterize the topology of the network by degree distribution p_k (fraction of nodes having degree k) and node dynamics by another probability distribution f_k (probability of removal of a node of degree k). The stability of networks is primarily measured in terms of a fraction of nodes called percolation threshold removal of which disintegrates the network into large number of small, disconnected components. We propose an analytical framework to measure the stability of peer-to-peer networks against peer churn and attack. During churn, nodes in the networks are removed randomly (random failure). During attack, important nodes are removed and the importance of a node is mainly characterized by its connectivity and bandwidth. Experimental validation of the theoretical results is done in two folds depending upon the generation of the peer-to-peer networks (simulating the network and using topological snapshot of Gnutella network). Next we report some of the significant observations of p2p networks in face of peer churn and attack.

Peer churn: **1.** It is important to observe that for the entire range of peer fractions, the percolation threshold is greater than 0.7 which implies that superpeer networks are quite robust against churn. **2.** Small fraction of superpeers in the network (specifically when it is below 5%) results in a sharp fall of percolation threshold, which shows that the vulnerability of the network drastically increases when the fraction of superpeers is low.

Attack: **3.** In networks with peer degree 1, 2 and 3, the removal of only a fraction of superpeers causes breakdown of the network. **4.** However as peer degree increases beyond 4, a fraction of peers is required to be removed even after removal of all the superpeers to dissolve the network.

Network emergence: We develop an analytical framework to explain the appearance of bimodal degree distribution in superpeer networks. The evolution of the network is mainly driven by the joining of incoming nodes through the bootstrapping protocol, departure of peers due to peer churn and rewiring of existing links. We review different bootstrapping protocols followed by the peer 'servents' like limewire, mutella etc and show that they can be easily modeled by various node attachment rules. Each incoming node joins the network with some finite bandwidth which restricts its maximum/cutoff degree. On the other hand, peers may leave the network randomly without any central coordination. This may result in the change in topological structure, increase in the network diameter and subsequently disintegration of the networks into small size components. In order to prevent network breakdown and to maintain the quality of service, rewiring of existing links takes place at the regular interval. Our framework shows that the interplay of finite bandwidth with node property and node dynamics plays key role in the emergence of bimodal distribution. We find that **1.** Increase in the cutoff degree reduces the amount of superpeers in the network. **2.** The analysis of the emerged network shows that the joining of the resourceful nodes initially increases the amount of superpeers in the network, but after a threshold level, the amount does not increase. **3.** A small churn results in a sharp reduction in the superpeer fraction in the network. More specifically, after a threshold amount of churn, the bimodality in the degree distribution disappears.

In stability analysis, we have assumed that the network is uncorrelated in nature. However, most of the real world technological networks like Internet, web network, peer-to-peer network (like Gnutella) exhibit degree-degree correlation in their network structure. We aim to incorporate this degree correlation in our analytical framework so that the theoretical results can able to provide exact results for real world networks also. Analyzing the emergence of superpeer networks is our ongoing research work. We believe that the rigorous analysis of the framework can reveal interesting conclusions regarding the topology of the emerging network. Subsequently, performing comparative analysis between the theoretical results with the real world network like Gnutella can provide us the 'self-organizing' local rules behind the emergence of the Gnutella network.



Chandan Karfa

Chandan Karfa received the B.Tech. degree in Information Technology from University of Kalyani, Kalyani, India in 2004 and the MS (by research) degree in Computer Science and Engineering from Indian Institute of Technology (IIT), Kharagpur, India in 2007. Since January 2008, he has been a research scholar in the Department of Computer Science and Engineering in IIT, Kharagpur. He received the Best Student Paper Award for his paper in ADCOM conference in 2007, Microsoft Research India PhD fellowship from Microsoft Research India in 2008, Innovative Student Projects Award (Master Level) from Indian National Academy of Engineering (INAE) in 2008, first prize in EDA contest in 22nd International Conference on VLSI Design 2009 and third prize in PhD poster contest in TechVista 2010 organized by Microsoft Research India. His current research interests include formal verification of circuits and systems and CAD for VLSI.

Supervisors: Prof. C. R. Mandal and Prof. D. Sarkar

Equivalence Checking in Embedded System Design Verification

Present day embedded systems synthesis consists in application of several sophisticated transformation techniques on the input 0 to improve its performance in terms of execution time, energy consumption, etc. Parallelizing code transformations are becoming increasingly important for multi-core/multiprocessor embedded systems. Sequential optimizing code transformations, human optimizations and transformations involved for design synthesis are also routinely applied. In this context, verification of the overall transformation is crucial for the reliability provided it meets the acceptable design cycle time. Formal verification is an attractive alternative to traditional methods of testing and simulation which, for embedded systems, tend to be expensive, time consuming, and hopelessly inadequate. While model checking and theorem proving based methods suit property verification, equivalence checking is the most natural choice for verification; it shows that executions depicted by the input 0 are equivalent to those depicted by the transformed behaviour. The objective of our research is to show the correctness of several transformations applied during embedded system synthesis primarily by equivalence checking techniques. In the following, transformation techniques along with our of to verify them are discussed.

The input behaviour is in the form of a sequential code. Several code motion techniques such as speculation, reverse speculation, branch balancing, conditional speculation, etc., may be applied at the preprocessing stage of embedded system synthesis. The input behaviours are transformed significantly due to these transformations. We have developed an FSM (finite state machines with datapath) model based method for checking equivalence between these sequential behaviours, i.e. the input behaviour and the transformed behaviour. Unlike many other reported techniques, this method is strong enough to handle both uniform and non-uniform types of code motions and the cases of control structure modifications of the original behaviours. Correctness and complexity of the method have been dealt with.

A high-level behaviour may be mapped to register transfer level (RTL) description during the synthesis process. The RTL consists of a description of the datapath net-list and a controller FSM. Towards establishing equivalence between

a high-level behaviour and its corresponding RTL behaviour, we have proposed a rewriting based method to extract the high-level behaviour from RTL description and then apply our FSM based equivalence checking method. Unlike many other reported techniques, our method is capable of validating pipelined and multicyle operations, if any, spanning over several states. The correctness and complexity of the presented method have been treated formally.

The functional specification, which relies on a single-threaded sequential code, is not easy to deploy on highly concurrent heterogeneous multi-processor systems. A parallel model of computation (MoC) may be used as the programming model. However, writing an application depicting concurrency is time consuming and error prone which conflicts with the low time-to-market requirement of the present day embedded systems. An automated tool that will convert a sequential code to its equivalent concurrent behaviour is employed. In the case of streaming applications, Kahn process network (KPN) model of computation is often used. We consider KPN for modelling the concurrent behaviours in this work. The proposed method consists of the following steps: The consistency of FIFO communications in the KPN behaviour is modelled as two properties the correctness of which can be checked using any theorem prover such as YIECS. The KPN behaviour is then linearized so that it can be modelled as an array data dependence graph (ADDG). Finally, an ADDG based method, proposed in this work, is used to establish the equivalence. This work is in progress.

The parallel process network model, obtained from the sequential behaviour in an automated way or manually, may not be suitable for the target architectural platform. In this case, it is necessary to manipulate the amount of concurrency in the functional model. The transformation techniques like, process splitting, channel merging, process clustering, and unfolding and skewing, may be used to control the concurrency in the KPN behaviour according to the architectural constraints. The verification task of this phase, therefore, is to show the equivalence between two KPN behaviours. We plan to extend our ADDG based method to handle those transformations.



Chester Rebeiro

Email: chester@cse.iitkgp.ernet.in

Chester Rebeiro received his MS(2009) degree in Computer Science and Engineering from IIT Madras and BE (1998) in Instrumentation and Electronics from Bangalore University. From July 1999 to May 2009 he worked for the Centre for Development of Advanced Computing. Since May 2009, he is a research scholar and a senior research fellow in the Department of Computer Science and Engineering, IIT Kharagpur. His research interests are in Cryptography, Computer Architecture, and VLSI.

Supervisor(s) : Prof. Debdeep Mukhopadhyay and Prof. Indranil Sengupta

Recent Trends in Side-Channel Attacks and Formal Modeling of Side-Channel Information Leakage

Most cryptographic algorithms used now-a-days are robust against a computationally bounded adversary. However in 1996 it was discovered that although the cryptographic algorithms are computationally secure, an implementation of the algorithm may leak secret information through covert channels such as power consumption, timing, and behavior in presence of faults. Such attacks are known as Side Channel Attacks. Many cryptographic algorithms in use today, like NIST's Advanced Encryption Standard (AES) have fallen prey to such forms of attack. A naive implementation would require just a small amount of side channel information to compromise the security of the entire system. It is therefore important to not only have robust cryptographic algorithms but to also have leakage proof implementations. However handcrafting such implementations and guaranteeing their security is difficult. This research focuses on developing tools and techniques that would provide users a platform to develop security infrastructure that is safe against classical cryptanalysis as well as side channel attacks.

Side channel attacks are broadly classified as active or passive attacks. Active attacks such as fault attacks tamper with the proper functioning of the system while passive attacks glean information from monitoring system's parameters such as power, timing, and radiation. The most common passive attacks are on VLSI chips and on devices that use cache memory. A simple power analysis of a VLSI chip can reveal the secret key with just a few power traces. Such attacks are easily preventable, but preventing the more deadly differential power analysis is not as straight forward. Encryptions that are run on devices that use cache memory have unequal encryption time. The variations in encryption time is dependent on the key. Thus simply monitoring the encryption timing is sufficient to obtain information about the secret key. The cheapest of all side channel attacks is the fault attack. An adversary can deliberately inject fault in the cryptographic device. By monitoring the cryptographic device under fault, the attacker can retrieve the secret key. The goal of the research is to develop a tool or to provide techniques that would notify the user about possible side channel leaks in the implementation. However the characteristics and attack algorithms used in each of the three side channels considered are different, making such a unified tool difficult to implement. As

the first step each of the three side channels are considered independently, and techniques would be developed to prevent leakages in each side channel.

When key recovery from side channel information is feasible, the security of the implemented cryptographic algorithm solely depends on the quantity of information about the key which can be extracted from the side channel measurements. The research involves building an information-theoretic model for side-channel leakage of cryptographic devices. In this, the device containing a secret key is modeled as a black box with a leakage function (which depends on the state transitions within the device) whose output is captured by an adversary as an observable parameter. This uses a theoretical concept of Mutual Information between the observed measurements and a hypothetical leakage to rank key guesses called Mutual Information Analysis (MIA). In other words, it studies the reduction in uncertainty on the guessed leakage function due the knowledge of a physical observable for a key hypothesis.

To counter side channel attacks targeting cache memory, the first step is to produce a generalization of the existing attacks and determine the cause of the leakages. The reason for cache attacks is due to the difference in the memory access time for a cache hit and a cache miss. It was found that the encryption time depends on the number of cache misses which in turn depends on the secret key. Another more subtle reason for cache attacks is in the distribution of cache misses with respect to the structure of the cipher. In modern processors cache misses are interleaved and parallelized therefore misses that occur together in a single round result in a lower miss penalty compared to cache misses occurring in different rounds where such accelerations cannot be done. To prevent such leakages, it is sufficient to protect only the first few rounds. These initial rounds are to be implemented such that the number of cache misses and the distribution of the misses should be uniform. This may require specifically designed encryption algorithms or tools which would alter the code at compile time to thwart leakages.

Fault attack is based on deliberate injection of fault in the cryptographic system. To thwart this attack the fault in the system needs to be detected. But it could be the case that while detecting fault, the countermeasures could open space for other kind of side channel attacks. So our current research in this area is to measure the robustness of existing countermeasures by exposing them to fault attack and other side channel attacks like cache attack and power attack. As a first phase of our research we are currently measuring the robustness of existing countermeasures on AES.

A unified tool to prevent all possible side channel attacks requires incorporation of all individual counter measures keeping in mind the performance overhead and the compatibility between the countermeasures.



Debi Prosad Dogra

Debi Prosad Dogra received a B.Tech degree in Computer Science from Haldia Institute of Technology, Haldia in 2001, and an M.Tech degree in Computer Science from Indian Institute of Technology Kanpur in 2003. From June 2003 till July 2006, he worked as a lecturer in Computer Science department of Haldia Institute of Technology, Haldia. From October 2006 till May 2007, he worked in the multimedia research team of Electronics & Telecommunication Research Institute, Daejeon, South Korea as a researcher. Since July 2007, he has been a research scholar in the department of Computer Science and Engineering at IIT Kharagpur. His research interests are in the areas of Image Segmentation and Video Surveillance.

Supervisors: Prof. A. K. Majumdar and Prof. S. Sural

A Computer Vision Based Approach for Conducting Hammersmith Infant Neurological Examination

Nowadays, application of computer vision to improve the efficacy of telemedicine and health-care systems is being considered as a good alternative. Advancement in image and video processing technology has accelerated its usage for betterment of daily livelihood of the mankind. For example, to help in the process of measuring exact postures of an infant while medical examinations are being conducted, a video based approach can be utilized. Certainly, it increases the efficiency and quality of experiments.

The work proposed in this article tries to address one such application where a computer vision based scheme can greatly increase the efficacy of the system. A widely used experiment that is carried out upon infants of less than two years of age is Hammersmith infant neurological examination. It is used for recording neurological development of the baby. While examination is going on, postures and reactions of the infant under consideration are recorded. An overall score that quantifies the neurological development index at the time of experiment is assigned to the infant. The experiment is carried out repetitively at different ages of the baby e.g. 3, 6, 9, 12 months and so on. Accurate measurement of posture, arm recoil, arm traction, leg recoil, leg traction, popliteal angle, head control, reflexes and reactions are parts of the experiment that are essential for estimating the overall neurological development index.

A computer vision based approach can be of great help in this regard. In this work, we propose a semi-automatic methodology to conduct the experiments. Some parts of the experiments can be carried out in a fully automatic manner while a semi-automatic scheme is proposed for remaining experiments. A complete description of the proposed scheme is as follows. The infant under consideration is placed on a mattress of white background. Cloths, hats or other objects for covering head are not used during experiments. This is done upon recommendation of physician. Two digital cameras used for capturing images / videos. Cameras are fixed such that one of it captures the top view of the infant (in supine posture) and the other one captures the front view (in sitting posture). We have used CIF frame format (frame size 352 X 288 pixels) for recording images. Illumination of the environment is kept

unchanged during the entire period of experiment. First frame is captured without any foreground object. It is used to eliminate background parts. A pixel-wise difference is calculated between current and background frames to obtain the foreground parts. Next, a morphological operator that performs erosion and dilation operations with a 3X3 structuring element is applied to remove any noise present. Even after morphological operations, however, some noise remains. To eliminate those noisy regions, contours of isolated regions are detected. The region with highest contour length is assumed to be the infant object. In order to cluster the foreground object into segments having similar characteristics, a clustering algorithm is applied. It is found that after initial classification many insignificant regions of tiny size are produced. Those segments are regarded as noise with respect to larger ones. To remove such insignificant segments, a region processing algorithm based on boundary sharing is proposed and used. Once initial classification of pixels is done, higher level of processing can be applied for object extraction.

Next, to extract the shape and boundary of the infant object, a model based approach is proposed. It is found through observations that different shapes of the infant can be parameterized using distinguishable features. For example, during posture estimation experiment, a method of skeletonization can be used to extract the skeleton of the baby and later it can be matched with some pre-existing models for assigning score. Images from the camera that is placed for capturing the top view of the infant at supine posture are used for this. Similarly, for arm and leg recall experiments, its output can be used to measure the final positions of the arms and legs after extension. In future, we have plan to use images from both cameras for extraction of experiment specific features e.g. location of head, position and dimensions of arms, legs and torso.

An image or video based semi-automatic scheme is always considered as a good alternative to many conventional approaches of conducting medical experiments. It can greatly increase the efficacy and accuracy of the experiments. In this work, we aim to propose a method for conducting the Hammersmith neurological examination. It is found that the tool can reduce the work load of the persons involved with such jobs. The problem is still a tough one and it is yet to receive enough attention. Our approach is convincing one and it can further be examined for a full scale implementation.



Dinesh Dash

Email: dineshd@cse.iitkgp.ernet.in

Dinesh Dash received his B.Sc. degree in Computer Science from University of Calcutta, Kolkata, India in 2000. He also received his M.Sc. in Computer and Information Science in 2002 and M.Tech. in Computer Science and Engineering in 2004 from the same institute. Dinesh worked as a lecturer at Asansol Engineering College, Asansol, West Bengal, India from 2004 to 2007. He is currently working towards his Ph.D. degree at the department of Computer Science and Engineering at Indian Institute of Technology Kharagpur. His research interests are in the applications of geometry in the field of sensor networks.

Supervisors: Prof. Arobinda Gupta and Prof. Arijit Bishnu (ISI Kolkata)

Geometric Problems in Wireless Sensor Networks

Wireless sensor networks have become an important area of research in recent years. A wireless sensor network (WSN) consists of a number of tiny devices with sensors to sense some parameters such as temperature, motion etc. A sensor has limited computation power, ability to communicate with nearby sensors, and limited battery power. Each sensor has a sensing range within which it can sense the parameter, and a communication range over which it can communicate with other sensors. The sensor nodes form an ad-hoc network capable of sending the sensed data to base stations for further processing. Sensor networks have been used in different applications such as habitat monitoring, intruder detection, target tracking etc. The challenges in designing protocols for wireless sensor networks usually stem from the limited resources (processing power, transmission bandwidth, memory, battery, etc.) available to each node. Most solutions requiring centralized processing or global information are likely to be wasteful for wireless sensor networks.

Geometric algorithms are used in different areas of sensor networks such as localization, topology control, coverage etc. Localization is the technique by which sensors can find their own coordinates in some coordinate system. Topology control is a technique by which sensors form and maintain network connectivity with certain desired properties. Coverage defines the quality of surveillance provided by the sensor network. We are working on the problem of coverage in sensor networks where a target region is to be covered by a set of sensors. There exist various definitions of coverage depending upon the application, such as area coverage, breach, support, barrier coverage etc. We are currently working on two problems related to coverage in sensor networks, as described next.

Let $X = \{s_1, s_2, \dots, s_N\}$ denote a set of N sensors, A denote a region of interest, $P = \{p_1, p_2, \dots, p_N\}$ denote the position of the sensors, and $d(p_i, p_j)$ denote the Euclidean distance between p_i and p_j . Given two points $i = (x_i, y_i), f = (x_f, y_f)$ in A , the support value ψ_p of a path P between i and f is the maximum distance of any point in P from its closest sensor. Formally, $\psi_p = \max_{p \in P} (\min_{s_i \in X} d(p, s_i))$. The maximal support path P_{max} between the pair of points (i, f) is a path whose support value is minimum among all paths between i and f . A path with a good support value is desirable

as it provides good coverage to an entity traveling along the path. In our first work, we focus on maintaining the support value of a path within some acceptable range on failure of sensors.

On failure of a sensor, the support maintenance algorithms try to move a small number of other sensors locally to establish a new path with an *acceptable* support value. We have presented two centralized approximation algorithms for support maintenance. The first algorithm allows the sensors to move in any arbitrary direction. The second algorithm allows the sensors to move only in one of the four directions - *north*, *south*, east, and west; the direction to move is chosen once and not changed after that. Support maintenance is actually a combination of two algorithms - an initialization algorithm that is called only once at the beginning to compute the maximal support path, and a maintenance algorithm that is called whenever the failure of a sensor is detected. The aim of the algorithm is to maintain a path with an acceptable support value. At each failure, the second algorithm establishes a new support path with support value close to that of the old support path by migrating a single sensor to a new location. Note that the new support path may not be the maximal support path corresponding to the final positions of the sensors. So far we have implemented the centralized as well as the distributed versions of the algorithms. We have shown that after the initialization phase, the support value of the maximal support path may change only if some sensor in the maximal support path fails. We have also shown that the new support value of the modified path after the first failure from the maximal support path is a constant factor approximation of the initial maximal support value. The amount of displacement of the sensor moved is also a constant factor approximation of initial support value. We have performed detailed experiments to evaluate the algorithm and shown that the actual performance of the algorithm is much better than the bounds given by the approximation factor.

In our second work we have defined a new type of coverage called *line coverage*. A line segment l is said to be *k-covered* if it is within the sensing range of at least k sensors. The input to this problem is a set of line segments $L = \{l_1, l_2, \dots, l_N\}$ and an initial deployment of a set of sensors $X = \{s_1, s_2, \dots, s_M\}$. The problem is to *k-cover* all the line segments by moving a minimum number of sensors to some new locations. To solve this problem, we are evaluating the minimum number of sensors required to *k-cover* the lines in L . If the number of sensors required is less than or equal to M , then there exists a rearrangement of the existing deployed sensors to *k-cover* the line segments in L , otherwise there does not exist such a rearrangement. Initially, we are focusing on finding the minimum number of sensors needed to *k-cover* the lines. We have shown that the problem of finding the minimum number of sensors to *k-cover* all the line segments in L is *NP-hard*. We have also designed a constant factor approximation algorithm for a special case of the problem when all the line segments in L are axis parallel. We are now working on the design and evaluation of good heuristic solutions to solve this problem.



Joydeep Chandra

Joydeep Chandra completed B.Tech (Computer Science and Engineering) from Haldia Institute of Technology, India in 2000 and M.Tech (Computer Science & Engineering) from Indian Institute of Technology, Kharagpur in 2002. From March 2002 till April 2008, he worked as lecturer in SLIET, Punjab and in Punjab Engineering College, Chandigarh. Since May 2008, he has been a research scholar at Indian Institute of Technology, Kharagpur. His research interests include p2p networks, distributed algorithms and complex networks. His website is <http://joydeep.chandra.googlepages.com>

Supervisor: Prof. Niloy Ganguly

Dissecting Peer-to-Peer Topologies Using Complex Network Theory

The enormous popularity of p2p applications - due to certain inherent benefits like, ability to share large contents directly from personal devices, enhanced scalability and robustness - has led to the formation of a pool of vast information and digital contents with entirely distributed set of entities and resources. However, to maintain the effectiveness of these networks, i.e. provide good quality of service, search performance, scalability and robustness, certain major issues need to be considered.

1. A major issue is the overlay or topology formation, the major objective being to obtain required contents with low bandwidth consumption. Since, there has been an unbridled growth in the p2p traffic and the ISP's are already facing huge problem of bandwidth and congestion, hence topology management has already become a major issue. An improper topology can increase p2p traffic at the ISP gateways. Moreover topologies also determine the robustness of the network, as an improper topology is susceptible to breakdown during network churn.
2. Another important issue that determines network efficiency is the search performance; because of the distributed nature of these networks, developing suitable search mechanisms is a big challenge. In p2p networks, many resources might be present that lack suitable computation power, hence we cannot implement search strategies that are computation intensive. Hence a major challenge is to implement simple and efficient search strategies.

Although, apart from these issues, there are certain other important issues like, service discovery, indexing and replication, and security that are being considered; however, the focus of my research is centered on improving performance like search, bandwidth wastage and bottleneck in heterogeneous p2p networks through better topology management schemes. In our approach, we would initially attempt to gain insight into the existing super-peer based network topologies like Gnutella, Kazaa etc. and analyze the effects of these topologies on the above stated parameters through suitable analytical models. Much of these models are influenced by the recent developments, in classical

physics, of network-theoretic models for analyzing the dynamics of large networks, typically termed as Complex Networks. These models help us to predict the behavior of the network under given conditions of growth and topology formation; further, these also provide directions to modify these topology formation mechanisms so as to improve the behavior of the network in terms of the above stated parameters. Hence, based on these findings we would attempt to propose suitable topology optimization mechanisms and validate their improvement in performance using simulations.

In this direction, we have initially studied the impact of topology on coverage and traffic redundancy in unstructured p2p networks that use flooding as the underlying search mechanism. In our work, based on models by Newman et al. [1] for finding neighbor distribution in large networks with given arbitrary degree distribution, we initially build up a basic analytical model to obtain probabilistic bounds on the network coverage of the peers in p2p networks that use TTL(2) (numeric value inside parenthesis represents the number of hops to search with) based search and query mechanisms. We limit our study to TTL(2) based networks, as search and query in popular networks like Gnutella mostly use TTL(2). However, our models and results can easily be generalized for TTL(n) searches as well. We observed that the basic model makes a simplified assumption that the underlying topology is tree-like; in contrast real networks contain certain cycle forming edges, which we referred as back and cross edges. We perform a rigorous analysis of bounds assuming the presence of back/cross edges. It is observed that the number of back/cross edges upon a node depend on its current degree. Based on this observation, we derive models to estimate the back/cross edge probabilities of the peers, with respect to their degrees, for random networks with any arbitrary degree distribution. The proposed refinement thus derives the TTL(2) network coverage of the peers with respect to their current degree. The accuracy of the models is validated on several types of networks, like Erdos Renyi networks, scale-free networks and also on a simulated Gnutella network, using extensive simulations. We further generalize the concept of back and cross edges for any TTL values and derive the TTL(n) network coverage of the peers in the network. The results indicate that the probability of occurrence of these edges increases enormously with increasing distance from the source nodes, which can result in huge traffic redundancy, thus questioning the effectiveness of larger TTL based search.

Based on these findings, we propose a modified bootstrap protocol for Gnutella networks, named HPC5 that avoids the formation of these back and cross edges. We analyze the performance of the proposed protocol and compare the same with an existing topology optimization mechanism named DCMP using simulations. The simulation results reflect that the proposed bootstrap protocol improves network coverage and reduces message complexity and traffic redundancy in Gnutella.

References:

1. M. E. Newman, S. H. Strogatz, and D. J. Watts. Random Graphs with Arbitrary Degree Distributions and Their Applications. *Phys Rev E Stat Nonlin Soft Matter Phys*, 64(2 Pt 2), August 2001.



Kamalesh Ghosh

Email: kghosh@cse.iitkgp.ernet.in

Kamalesh Ghosh received a B.Tech. (Hons) degree in Computer Science and Engineering from IIT Kharagpur in 1998. From July 1998 to April 1999 he worked as a software engineer with Wipro Infotech Ltd. (Bangalore) on e-commerce products. From April 1999 to Dec 2000, he worked as a senior software engineer in Delsoft India Pvt. Ltd. (Noida) on Electronic Design Automation (EDA) software. From Jan 2001 to Oct 2004 he worked as senior R&D engineer at Synopsys Inc. (Marlboro, MA) on verification tools for VLSI design. From Nov 2004 to Nov 2007 he worked at Synopsys India Pvt. Ltd. (Bangalore) as senior R&D Engineer, continuing in the same area of work. From Dec. 2007 till now, he has been working as a Research Consultant in the department of Computer Science and Engineering at IIT Kharagpur, pursuing a Ph.D. degree simultaneously. His research interests are in the area of Artificial Intelligence and Formal Verification with particular focus on application to component based design of safety critical real-time systems.

Supervisor: Prof. Pallab Dasgupta

Formal Methods for Top Down Component Based Development

Component based Software Engineering (CBSE) is a very popular paradigm in modern software engineering. The CBSE approach focuses on building software systems with commercial-off-the-shelf (COTS) components or existing in-house components rather than ground-up development. When safety critical systems with real-time requirements (e.g. automotive) are built using this paradigm, sources of failures can be many. For example - the timing and logical properties of the built system are inherently difficult to predict or verify. Our work is focused on finding novel techniques that may help in closing some of these sources of failure.

To give a proper definition to our task, we visualize three abstract layers across which the design and implementation of the system is distributed. The topmost layer is named the Feature Layer in which the requirements of the built system are captured from a user's perspective. This layer is the most idyllic view of the system which will just list desirable features and have no connection to lower level concerns. The second layer, named Interaction Layer, is a cluster of various "subsystems" which coalesce together to build up the system. Each "subsystem" may be thought of as a component in our CBSD paradigm, which is being bought as a COTS component or developed independently in-house by the manufacturer, e.g. the braking subsystem or the powertrain subsystem for a car. Though this layer is still not giving a complete picture of the working of the whole system, it is more grounded towards reality and detailed. The lowermost layer, called the Component Layer, is where the real implementation is captured. This layer takes into account all the implementation details - like the actual hardware platform and physical interconnection -- into account. This 3-layer visualization mimics the phases in the design of a real-life system quite realistically. Based on the above framework, we define some point problems which are inherently interesting, challenging and potentially useful in producing a whole solution eventually.

In our first problem, the interaction layer specifications are formally written as sets of preconditions and postconditions. Each precondition-postcondition pair is called an action and either defines what the controller must do when the preconditions hold or defines what the environment (driver, road etc.) may do if it chooses to. In the former case the actions are called control actions while in the latter case we call them environment actions. Thus our formalism includes the operational environment and control specification of the system as its core elements. The feature layer is simply modeled (for now) as a set of logical statements which indicate desired properties (checks) for the system. The control should never allow any of these to be violated (intermittent violations are allowed, but the control should never allow the system to sustain such a violated state). We model the environment and control as two adversaries in a game-like scenario. The environment makes moves to violate a property representing a vehicle feature requirement, while the control interrupts at every move of the environment and executes pre-specified actions. The property is verified if the environment has no winning strategy. This model allows us to do a logical evaluation of the software control logic at a stage when few low level details are available. The benefit of this analysis is that we may detect "logic bombs" at a very early stage of design.

In our second problem, building up on the same formalism above we aim to catch contradictions or inconsistencies in the specification through automatic detection of loops consisting of control actions. Loops in the high level specification of a control naturally arouse suspicion as it can be indicative of contradictions. We have demonstrated this point through examples in our related paper.

Further building up on the work done till now we are looking at methods to expand the semantics of our formalism to consider loops in a more realistic manner. We assumed in the above problem that loops in control are indicative of an inconsistency. In reality, this need not be the case. For example, many automotive features may require components to continuously interact with each other till a particular event happens in the environment. This can naturally lead to loops in the control action definitions. We are exploring further enhancements to our existing semantics in order to consider these loops in a smart and realistic way. This work requires us to define new semantics and devise new algorithms.

Specifications for real-time reactive systems often need to refer to numerical value of physical quantities such as speed, acceleration etc. Any formalism without this basic expressive power may be considered too limited for practical use. However, allowing for expressions with numerical variables under standard operations like addition, multiplication etc. often causes the verification problem to become undecidable. Our future research will explore limited enhancements in expressive power in the numeric domain to find a good tradeoff between expressive power and ease of verification.

This research is supported by a grant from General Motors under the GM-IITKGP Collaborative Research Lab.



Maunendra Sankar Desarkar

Maunendra Sankar Desarkar received B.E. degree in Computer Science and Engineering from the University of Burdwan in 2004, and M.Tech. degree in Computer Science and Engineering from Indian Institute of Technology Kanpur in 2006. From July 2006 till July 2008, he worked for Sybase India Pvt. Ltd. as a Software Developer. Since July 2008, he has been a research scholar in the department of Computer Science and Engineering in IIT Kharagpur. He was awarded the Microsoft Research India PhD Fellowship in 2009. His research interests are in the areas of Data Mining and Information Retrieval.

Supervisor: Prof. Sudeshna Sarkar

Aggregation of Rankings and Ratings in Information Retrieval Applications

Several problems in web domain require aggregation of opinions obtained from multiple sources. The opinions can be human or system generated. System generated opinions can be outputs for the same problem obtained by employing different algorithms, for example, rankings of documents by search engines, scoring of documents or objects based on multiple features etc. Human generated opinions include ratings assigned to movies/news/products by users, free text reviews in blogs, forum etc. We tried to focus on the problems of meta-search and collaborative recommendation, two applications which require aggregation of opinions.

The number of pages in the web has increased manifold in the recent years. The contents have become richer, and also the linkage pattern between pages has changed a lot. The changes are posing major challenges to the Information Retrieval problems that work on web data. For web search, which is one of the most widely used Information Retrieval applications, the problem is more severe as it is inherently difficult for the search engines to capture the user intent and the context of the documents. Under the circumstances, it might be helpful to combine the outputs of multiple search engines for getting the final output to be displayed to the user.

Meta-search engines are special kind of search engines that take help of other search engines for producing search results. Upon receiving a query from a user, it passes the query to multiple other search engines and obtains the search results from them. The search results are given in the form of ranked list of documents. The meta-search engine combines these ranked lists to generate an aggregated ranked list. This ranked list is then displayed to the user. The first problem that we tried to address is the Rank aggregation problem, where the task is to aggregate the ranked lists obtained from multiple different sources. The problem has applications in text mining as well, where documents are scored based on multiple features or criteria. For each criterion, the documents in consideration can be ordered in the form of a list. The final ordering of the documents can be constructed by combining the lists obtained for the individual criteria or features.

We propose a two-phase approach for solving the rank aggregation problem. Each ranklist is represented by a preference graph. The first phase of our algorithm assigns weights to the input ranklists. We want to assign higher weights to better rankers, so that poor rankers do not have much influence on the aggregated result. The input preference graphs are combined to produce a weighted aggregate preference graph. The second phase of the algorithm induces a linear ordering from this aggregate preference graph. We have tested our algorithm on synthetic data. Currently we are in the process of evaluating the algorithm on a benchmark dataset.

Another problem that we are simultaneously working on is related to collaborative recommender systems. The web has evolved into a platform where users can share their thoughts and discuss ideas. Introduction of e-commerce services and sharing of customer experiences through product review sites or even personal blogs are providing the potential customers with a huge pool of information. Since the judgments are provided as a conscious effort to let the other users know about the experiences of the customer/reviewer/author, these are extremely valuable sources of information. However, the sheer size of such data often makes it difficult for an individual to draw any meaningful conclusion out of it. Systems able to aggregate such opinions can be immensely beneficial to the users.

Recommender systems aim to solve this problem of information overload by presenting the users a small subset of products that he might be interested in. Effective recommendation helps both sides of the business, the customer as well as the merchant, and has emerged as an important research area. Users in recommender systems often express their opinions about different items by rating them on a fixed rating scale. The rating scale can have two categories - like/dislike, or multiple categories where higher categories denote higher level of user satisfaction. Primary task of such recommender systems is to predict the rating that a user would give to an item he has not rated yet. Based on the assumption that users with similar tastes rate items similarly, collaborative recommendation systems first find a group of users having similar interests. Opinions of the users from that group are used to predict the unknown rating. We have proposed a memory-based collaborative filtering approach that uses preference relations between items instead of absolute ratings.

Use of preference relations allows the rating of an item to be influenced by other items. This is in contrast with weighted-average approaches of the existing techniques. This is certainly desirable, especially when the data is sparse. We carried out experiments with sparse data to verify this claim. The step for finding similar users requires assigning similarity weights to the users. Possible extensions of the work may be to update the similarity weights dynamically as new ratings are added to the rating database, or to try different similarity measures for determining the similarity weights.



Plaban Kumar Bhowmick

Plaban Kumar Bhowmick received BTech degree in computer science and engineering from University of Calcutta in 2002 and the MS degree in computer science from Indian Institute of Technology, Kharagpur in 2006. He is currently a research scholar in the Department of Computer Science & Engineering , IIT Kharagpur. His research interests include natural language processing, information retrieval, information and communication technology.

Supervisors: Prof. Anupam Basu and Prof. Pabitra Mitra

Reader Perspective Emotion Analysis of Text in Multi-Label Classification Framework

The Internet has been one of the primary media for information dissemination after the advent of World Wide Web. The advent of new technologies makes way of new interaction possibilities and provides people to perform different social activities on platforms like blog, chat, social network, news etc. As compared to traditional keyword based or topical access to the information, social interactions require the information to be analyzed in social and humanistic dimensions like emotion, sentiment, attitude, belief etc.

Opinion mining or sentiment analysis focuses on investigating the views of the users towards a particular entity. This task judges the entity in the dimension of positivity or negativity, i.e., whether a particular movie is liked by the users or not. In contrast to sentiment analysis, emotion recognition goes beyond positive-negative dimension to discrete emotion categories like happiness, disgust etc.

Expression or change of behavior is the most visible and prominent clues for recognizing emotion. Facial expressions, speech expressions have widely been used in detecting emotion. Emotion is not a linguistic entity. However, language is one of the most common modes for expressing emotion whether it is day-to-day speech communications (spoken language) or published communications (written language). Emotion can be studied from two perspectives.

- a) From the writer/speaker perspective, where we need to understand the emotion that the writer/speaker intended to communicate and
- b) from the reader's perspective, where we try to identify the emotion that is triggered in a reader in response to a language stimulus.

In this work, we intend to perform sentence level emotion recognition from a reader's perspective. The challenges included in this task are as follows.

- **Data Preparation:** Appropriate data set for emotion analysis is to be collected and annotated by human judges with proper annotation scheme.

- **Corpus Reliability:** As Emotion is a subjective entity, the emotion tagging may vary with judges. So, annotation by multiple judges may be required to test the reliability of the corpus before using it as a gold standard in classification task.
- **Multi-Label Characteristics:** On encountering a sentence as stimulus, a blend of multiple emotions can be triggered in a reader. For example, the following sentence may evoke fear and sad emotion in readers mind.

Militant attack kills over 30 persons in Nigeria.

Hence, a sentence may be labeled with multiple emotions from the reader's perspective.

- **Feature:** Since research on emotion elicitation from text is in its infancy, the appropriate feature set required for emotion elicitation needs investigation.
- **Sparseness of features:** While emotion elicitation from a discourse or paragraph may provide larger number of cues as features, the number of features available from a single sentence is less and hence the feature space becomes sparse.

The emotion text data collected by us consists of 1350 sentences extracted from Times of India news paper archive. The sentences were collected from headlines as well as from the bodies of articles belonging to political, social, sports and entertainment domain. The corpus has been annotated by five human judges. The annotation scheme considers fuzzy annotation. Two different measures have been formulated in order to compute agreement among the annotators. The crisp reliability measure considers crisp belongingness of a sentence in an emotion category whereas the fuzzy agreement measure considers fuzzy belongingness. Gold standard data sets have been prepared by aggregating the views of all the annotators. As each sentence may evoke multiple emotions simultaneously, multi-label classification frameworks have been used to classify sentences into emotion categories. Multi-label k Nearest Neighbor and Random k Label Set algorithms have been utilized to develop classification models. An extensive study of features for reader emotion analysis has been provided. Four different types of features have been considered, namely, word occurrence, polarity (subject, object and verb phrase), semantic frame, emotion elicitation context. The same set of features has been considered for developing fuzzy classification model with Fuzzy k Nearest Neighbor algorithm.



Priyankar Ghosh

Priyankar Ghosh received B.E. degree in Computer Science & Engineering from Jadavpur University, Kolkata, in 2003 and M.Tech degree in Computer Science & Engineering from Indian Institute of Technology Kharagpur, in 2006. He also has industry experience of approximate two years. Since April 2007, he has been a research scholar in the Department of Computer Science & Engineering in Indian Institute of Technology Kharagpur. His research interests are in the areas of Verification, Artificial Intelligence and Knowledge Representations and Interoperability among them.

Supervisors: Prof. Pallab Dasgupta and Prof. P. P. Chakrabarti

Formal Methods for Planning and Verification of Integrated Semantic Web Services

With the recent advances in internet technology, Service Oriented Architectures (SOA) have gained widespread acceptance. Typically web services that implement SOA, represent functionalities that are offered by some organizations in the web. These functionalities can be accessed through internet by some client, which may be an individual or an organization. Web services resemble remote procedure calls, which are accessed using HTTP/HTTPS protocol.

The web service requester does not need to know about any implementation details of the web service provider. Therefore the interoperability among different organizations increases greatly. Web services are published, described, and accessed by certain machine processable descriptions developed on top of XML. Moreover existing web services can be combined in a loosely coupled fashion to develop complex applications.

Semantic web is an ongoing extension of traditional web where the semantics of the service is defined. The main goal of semantic web is to enable the machine to interpret this information. Semantic web services are a component of the semantic web activity where machine-readable markups are used to describe a service. The objective of semantic web services is to automate the discovery and invocation of the services.

Since the origin of the World Wide Web, the development and growth of web services has taken place typically in an uncoordinated and unstructured way. Consequently the protocols followed by different web services are vastly different, not only in terms of the protocol structure but also in terms of the semantic interpretation of the data they exchange. This makes the task of developing applications which automatically interact with multiple web services, a significantly challenging task.

In the current work we study modeling techniques and investigate the usage of the high level semantic models for solving the following problems.

- **Usage of High Level Model in Protocol Verification:** Web service providers typically publish a high level model of the service to describe the behavior of the web service. Typically these models are written in English language and may have some graphical representation as well. In this work our goal is to formalize the model, and generate a set of test cases using the model in order to verify the correctness of the implementation. These test cases will include positive as well as negative test cases. This published model is also used during the integration with the client. During the integration with the server, the client side writes test cases to check the protocol compliance. Diagnosing the reason of the mismatch in message exchange between the client and the server plays a crucial role in debugging the client applications. Our goal is to develop and formalize a debug mode based testing methodology which will assist the client to detect reason of interaction failure during the integration with server and provide useful information regarding the test cases developed by the client.
- **Modeling Semantic Information Exchange and Detecting Conflicts:** Interaction between the client and the server may also fail due to the difference in the interpretation of the exchanged data. The semantics of data play a major role in semantic web services which goes beyond simple type checking. Therefore the protocol that defines the interaction has to be verified in order to check the presence of semantic conflict. It is possible that the knowledge base of the server has some conflict with the knowledge base of the client, but the protocol does not sensitize the conflict.
- **Planning Based Approach to Compose Web Services:** It is quite common to use multiple web services in order to achieve a goal. Since the overall goal may be achieved only when these web services are invoked in a particular order, it is often needed to undo the effect of one service. For example hotel reservation may be cancelled due to the unavailability of flight booking. However these cancellations may incur penalty. Moreover alternative web services may be available for the same goal. For a given goal, the objective is to find out a schedule of invoking the web services so that the penalty in this schedule is minimized.

This research is partially supported by a research grant from Google.



Rajarshi Pal

***Rajarshi Pal** received a B. Tech. degree in Information Technology from Kalyani Govt. Engineering College in 2004, and an M. Tech. degree in Software Engineering from Jadavpur University in 2006. Since July 2006, he has been a research scholar in the Department of Computer Science & Engineering in IIT Kharagpur. His research interest is in the area of Computer Vision.*

Supervisors: Prof. Jayanta Mukhopadhyay and Prof. Pabitra Mitra

Modeling Visual Attention for Image

Machine vision tasks enable the computer to understand what it "sees". Hence, any machine vision based task that works with images needs to analyze the captured image. Often the next step of processing is determined by this analysis or understanding of the image. As still now computers are much slower compared to primate brain, they take a lot of time to analyze the captured media. Instead of analyzing the whole image, if analysis from some selected portions of it gives comparable result, then computational time is saved. The concern is which portions of the image will be analyzed and which portions will be ignored. Obviously, in order to be not distorted from the analytical result, important portions of the image have to be analyzed. Here comes the urge of deriving a computational model that conveys to the machine about the important portions of the image. In other words, the model will direct the machine to attend the selected portions, like the primates do.

Without doubt, primates are remarkably faster to comprehend and analyze visual scenes in real time. It is observed that in a time interval that is short enough to comprehend the whole scene, primates tend to attend some selected areas or objects that can be easily distinguished (salient) from a cloud of other objects. In order to perform computer vision based tasks faster, researchers in this field are trying to mimic primate vision.

A complex network based approach is adopted to determine the visually salient locations in the image. A network is constructed where nodes represent accumulation of similar pixels and the dissimilarity between any pair of such accumulations is encoded as edge-weight between the corresponding nodes. Dissimilarity is in terms of features modulated by their positional proximity. Modulation by positional proximity ensures that difference with neighboring locations gets more weightage. Such networks are constructed over multiple features (color, intensity and orientation) across multiple scale representation of the image. Certain location is salient if it is dissimilar from all other locations and specially, with its surroundings. The network, termed as ViSaNet (Visual Saliency Network), constructed here is a suitable choice for determining salient locations as it combines both local and global conspicuity of a location. Incorporation of degree centrality analysis with this type of network suggests that a centrally situated node belongs to a salient location. Potential of other node centrality measures (eigenvector, closeness, and betweenness) to model saliency of an image is also studied.

Selection of appropriate features is an important task in computer vision problems. It reduces the computational burden and improves the end result. Important features for visual saliency models are identified by estimating the amount of errors contributed by them. Analysis for monochrome and color images is carried out separately. Experimental results suggest that selection of appropriate features improves the model's performance. Principal component analysis (PCA) on selected features is also performed. Usage of only the first principal component as potential feature in degree centrality based saliency computation performs better than well-known saliency models.

Evaluation is a key part while proposing a new model. To evaluate models of visual saliency, one needs to compare the model's output with salient locations in an image. An approach to find out the salient locations, i.e., ground truth for experiments with visual saliency models, is derived purely based on human hand-eye coordination. It is found that the proposed technique can be an alternative to costly human pupil-tracking based systems. Moreover, an evaluation metric is also proposed that suits the necessity of the saliency models.

Red, orange and yellow are termed as warm colors. Role of warm colors on visual saliency is studied. Distributions of chromatic features (hue and saturation) are found to be different for warm colors that draw attention and those that do not. It is also observed that likelihood of drawing attention by a warm color depends on both of its hue and saturation component. Interestingly, this dependency is not related to the absolute hue and saturation, but relative to these chromatic components of other warm color pixels. Warm colors with hue relatively closer to red and/or with high saturation are more likely to guide attention.

So far, various issues related to modeling of visual saliency have been discussed. Beside the works that tune up model of saliency step-by-step towards perfection, we have also worked in a few application areas of visual saliency.

Image downsizing to show it in a small display reduces recognizability of the image. The standard practice is to crop around the important locations and display the cropped portion only. A cropping algorithm has been proposed that uses localized thresholding to retain the overall subject matter of the image. This method also does aspect ratio adjustments to prevent any distortion due to mismatch in aspect ratio of the cropped portion and that of the display.

Continuous inventions of modern cameras are making the act of good photography easier. We have proposed a methodology that makes a camera even smarter by converting the objects of interest to the photographer more salient. This results in maximization of viewers' attention on intended objects. The captured photo can be taken as a media of communicating photographer's mind to the potential viewers. In that sense, the proposed methodology reduces the communication gap between them.



Rajiv Ranjan Suman

***Rajiv Ranjan Suman** received a B.E. degree in Computer Science & Engineering from Birsa Institute of Technology, Sindri, Dhanbad (Jharkhand) in 1991, and an M.Tech. degree in Computer and Information Technology from IIT Kharagpur in 2002. From Mar 1992 to Dec 1995, he worked in BIT Sindri as a part-time Lecturer. From Jan 1996 to till date, he has been working as a Lecturer in the department of Computer Science & Engineering at NIT Jamshedpur. In July 2007, he joined the department of Computer Science & Engineering, IIT Kharagpur, as a research scholar as a sponsored candidate under QIP scheme. His research interests are in the areas of Software Engineering.*

Supervisor: Prof. R. Mall

Construction of State Model of Software Components

We propose a novel black-box approach to reverse engineer state models of software components. We assume that in different states, a component supports different subsets of its services and that the state of the component changes solely due to invocation of its services. To construct the state model of a component, we track the changes (if any) to its supported services that occur after invoking various services. Case studies carried out by us show that our approach generates state models with sufficient accuracy and completeness for components with services that either require no input data parameters or require parameters with small set of values.

In component-based software development paradigm, a large software is built by assembling pre-built and independently developed “plug and play” type of executable units, called software components. Components are usually integrated to applications through an application program interface (API). Only a brief description of the functionality of a component is provided by the component vendor. Developers of the component-based software have to rely on vendor's capability and inadequate documentation available regarding correctness of the functionality and quality of the components. However, developers of critical applications cannot risk using components of incorrect functionality and they need to ensure that the components are trustworthy and would function as per the expectation. In addition to validating the functional behavior, dynamic behavior of the components need to be validated.

As the source code of components is not available, correctness and limitations of component functionality may be analyzed using two major methods of black-box understanding: binary (object code based) reverse engineering and interface probing [1]. Binary reverse engineering is an effective way to understand the design and overall functionality of a component. However, this requires a large number of test cases to be created and analyzed. Interface probing reveals some of the component properties which are often inadequate for the component developers.

State models (FSMs, statecharts, etc) of objects in object-oriented systems is a behavioral model that depicts the different states that the object may assume and transitions among the states that may occur in response to the stimuli received from its environment. A component is a generic term and is often designed and implemented as a single class

or a collection of classes or no class at all (in case of procedural implementation), or it may even integrate many smaller components. In spite of the apparent diversity in component implementations, as far as state modeling is concerned, a component in the component paradigm can be considered analogous to an object in the object paradigm since both a component and an object can be considered as black boxes that store some data and provide some externally visible behavior. Externally, the state model of a component is visualized in terms of the state-based behavior of the component as a whole rather than in terms of the individual objects that the component may be composed of.

In component paradigm, component in an application can effortlessly be replaced any time by another functionally equivalent component. After every such change to a component of a critical application, regression testing of the application needs to be carried out to ensure that the various features continue to work satisfactorily even after component upgradation. Selection of regression test cases for component-based software is considered a challenging research problem. [2]. However, components are usually not accompanied with their state models. In the absence of a state model, it is difficult to test the state behavior of a component.

State-based bugs are difficult to detect using traditional testing techniques [3]. A system might behave correctly to a user's requests in only some of the states but not in other states. It is also possible that the system may not transit to some required state even when all necessary conditions are satisfied (missing transitions) or may have improper transitions (sneak transitions) to certain states. State-based software testing has therefore been accepted as a crucial type of testing that can help detect such insidious bugs. State models form an important basis for state-based testing in the component paradigm [4]. State coverage and transition coverage are two popular state-based testing techniques [3].

Components are increasingly being used to build embedded systems, distributed control applications, and several types of real-time systems. These applications mandate ensuring high degrees of reliability, safety, and security. In this light, state model-based testing assumes importance. Besides its use in testing, the extracted state model of a component has several other applications as well. These include understanding the state-based behavior of a component and re-engineering of a component to meet new requirements or constraints. A state model can also be used to estimate the complexity and effort needed for state-based testing, as well as to estimate the reliability of a component.

We represent the state models of components as FSMs as they are easy to use, very intuitive and popular. However, FSM lacks hierarchy and concurrency and suffer from state explosion problem. Statechart [5] was proposed to tackle these problems of FSMs. During construction of state models, software designers often end up developing the FSM models of design elements rather than their statechart models. Further, during the reverse engineering of legacy code, an FSM model is naturally constructed rather than a statechart model. Therefore, it is often required to convert FSM based state models to statecharts. However, very few research work are available in literature on conversion of FSM to statecharts. Methods discussed in these works are either incomplete to introduce hierarchy and concurrency to the FSM or they are inefficient for handling large FSMs. We propose efficient methods for converting FSM models to statechart models.

References

1. **Korel, B.:** "Black-Box Understanding of COTS Components" Proceedings of the 7th international Workshop on Program Comprehension (May 05 - 07, 1999). IWPC. IEEE Computer Society, **Washington, DC**, 92.
2. **Gao, J. Z., Tsao, H. S. J., and Wu, Y.:** "Testing and Quality Assurance for Component-based Software", Artech House publication, Norwood, **USA**, 2003.
3. **Binder, R. V.:** "Testing Object-oriented Systems: Models, Patterns, and Tools", Addison Wesley Longman Publishing Co, Inc, **Boston, MA**, 1999.
4. **Gallagher, L., Offutt, J., and Cincotta, A.:** "Integration testing of object-oriented components using finite state machines", Software Testing, Verification and Reliability, Vol. 16(4), Jan 2006, pp. 215 - 266.
5. **Harel, D.:** "Statecharts: A visual formalism for complex systems", Science of Computer Programming, 1987, Vol. 8(3), pp. 231-274.



Sanjay Chatterji

Email: schatt@cse.iitkgp.ernet.in

Sanjay Chatterji earned B. Tech. degree in Computer Science and Engineering from the Haldia Institute of Technology in 2003, and M. E. degree in Computer Science and Technology from B.E. College, Shibpur in 2005. He worked as a lecturer in CSE Department of HIT, Haldia for 1 year and in CSE Department of KNSIT, Bangalore for 1 year. Since January 2008, he has been a research scholar at the department of Computer Science and Engineering in IIT Kharagpur. His research interests are in the areas of Machine Translation and Natural Language Processing.

Supervisors : Prof. Sudeshna Sarkar and Prof. Anupam Basu

Hybrid Approaches to Machine Translation and Parsing

Machine translation is the application of computers to the task of translating texts from one natural language to another. In this process the meaning of the source language must be preserved in the generated text in the target language. The translation process may be stated as decoding the meaning of the source text and re-encoding this meaning in the target language. There are mainly two ways of performing the machine translation - rule based machine translation (RBMT) and statistical machine translation (SMT). RBMT relies on built-in linguistic rules and resources. Developing a RBMT involves a great deal of time and linguistic expertise. SMT on the other hand provides good quality translation when large and good quality parallel corpus is available. It creates statistical model from the parallel corpus. But when parallel corpus are not available or are only available in limited quantities, one may go for a combined approach making use of simple rules and supplemented by small parallel corpora.. Recent works on SMT have led to significant progress in coverage and quality of the translation systems but the amount of work involving translation into Indian languages, like Bengali or Hindi, is quite limited.

We have attempted to combine different types of knowledge and developed a hybrid system for the Bengali-Hindi language pair. The overall translation quality is improved by exploiting explicit linguistic knowledge contained in the rules and resources of the RBMT system and implicit knowledge that can be extracted from the SMT system. We have used language resources such as dictionary, suffix list, rules etc to improve the SMT output quality. The open source decoder called Moses is used to get the baseline output from the parallel corpus. We have used a modification of BLEU score methodology to consider concepts rather than words in evaluation process. We have restricted our experiments for Bengali to Hindi machine translation to the news domain. We have only a 12,000 sentences parallel corpus and a bilingual (Bengali-Hindi) concept dictionary of size 20,000 concepts. We have also crawled some 500K word corpora of Bengali and Hindi languages each. Using these resources and developing the rules we got a Bengali to Hindi Hybrid Machine Translation System with the BLEU score of 22%.

Recently some work has been done on a transfer based Bengali to Hindi machine translation system. The Bengali to Hindi ILMT project sponsored by MCIT, Govt. of India is one of the most recent transfer based system which

gives the BLEU Score of 5%. It gives a baseline word sequence of target language for a given source language word sequence. But the target word sequence needs to be corrected to some extent to get a meaningful target sentence. For that purpose we are using a language model based approach. The target language sentence is generated in steps. The sequence of translated words is chosen so as to maximize the probability of the sequence according to the language model of the target language.

This is a step by step decoder. A word graph based decoder can be more effective. The word graph or the word automata is a weighted graph which has one start and one end states. It represents all possible translations for a given source language word sequence. This decoder will use a word hypotheses graph for each source language sentence as an efficient search space representation and find out the most probable word sequence using a dynamic programming approach. Viterbi algorithm may be useful in this context. This decoder will use the Synset dictionary to find the most probable target language word sequence for a particular source language word sequence with the help of the n-grams (n previous words) of a previously created corpus.

Finally our target is to use the parsed corpus to reorder the phrases of the target sentence. Bengali and Hindi have relatively similar grammatical structure. Both have complex morphology and relatively free word order. But the translations of the other language pairs like Bengali-English need the phrase reordering. The Parts of speech, chunking, parse and all other morphological features are very important to get the good machine translation output, at least for morphologically rich languages like Bengali and Hindi. So these features will be used in the decoding. The reverse is also true. That is translation also can help in parsing. So, if we have a Hindi parser and a good Hindi to Bengali translator then we can achieve a good Bengali parser. I shall also try this approach in my research work.



Santosh Ghosh

Santosh Ghosh received his B.Tech. degree in Computer Science and Engineering from Haldia Institute of Technology, Haldia 721302, India in 2002. He received his M.S. degree in Computer Science and Engineering from Indian Institute of Technology, Kharagpur, India in 2008. He is currently working towards his Ph.D. degree in the department of Computer Science and Engineering at Indian Institute of Technology Kharagpur. His research interests include cryptography and network security, hardware implementation of cryptosystems, and side-channel analysis.

Supervisors : Prof. Debdeep Mukhopadhyay and Prof. Dipanwita Roychowdhury

Design and Analysis of Pairing Based Cryptographic Hardware

Bilinear pairing is a one-way function defined on elliptic or hyperelliptic curve group. It has attained utmost importance in the field of public key cryptography. This area of research is known as pairing based cryptography. It has a wide application area. Most of the identity based encryption and signature schemes belong to pairing based cryptography. In practice, pairing based cryptographic protocols are mainly used in identity aware and ubiquitous computing devices. The security protocol in these handheld devices either runs on a dedicated VLSI chip or on a low-end general purpose embedded processor.

The name bilinear pairing indicates that it takes a pair of vectors as input and returns a number, and it performs linear transformation on each of its variables. For example, the dot product of vectors is a bilinear pairing. Similarly, for cryptographic applications the bilinear pairing (or pairing) operations are defined on elliptic or hyperelliptic curves. Pairing is a mapping $G_1 \times G_2 \rightarrow G_3$, where G_1 is a curve group on some field F_q , G_2 is another curve group on the lowest extension field F_{q^k} , and G_3 is a subgroup of the multiplicative group of F_{q^k} . Let, P be a member of G_1 with order l . The mapping satisfies following properties:

- Non-degeneracy : For each $P \neq O$ there exist $Q \in E(F_{q^k})[l]$ such that $e_l(P, Q) \neq 1$.
- Bilinearity : For any integer n , $e_l([n]P, Q) = e_l(P, [n]Q) = e_l(P, Q)^n$ for all $P \in E(F_q)[l]$ and $Q \in E(F_{q^k})[l]$.
- It is efficiently computable.

Pairing computation and elliptic curve scalar multiplication (ECSM) are two major operations that are essential for pairing based cryptography. Both of them are inherently computationally complex and hence their implementation on target devices is a real challenge on all platforms.

Bilinear pairings such as Weil pairing and Tate pairing in cryptography were earlier used for cryptanalysis. It is first used for reducing elliptic and hyperelliptic curve discrete logarithms to logarithms in finite field. The positive applications of bilinear pairing in cryptography for constructing cryptographic primitives have been found in recent years. Joux,

in 2000, introduced the application of bilinear pairing for constructing cryptographic protocols. The paper proposed a three-party one-round key agreement protocol using Weil pairing. Subsequently, Boneh and Franklin presented the first fully functional, efficient, and provably secure identity-based encryption scheme using the properties of bilinear pairing on elliptic curves.

However, we choose the current area of research by motivating from the following facts. Due to the bilinear pairing, the decisional Diffie-Hellman (DDH) problem becomes easy. Thus the existing public key protocols that are based on the difficulty to solve DDH problem becomes weaker. As an alternative, the protocols are developed on bilinear pairings where the security is based on the difficulty to solve bilinear Diffie-Hellman (BDH) problem. And till date there are no weakness found on above problems. Literature says that surprisingly sufficient works have been done on the development of pairing based cryptographic schemes, whereas very few works are there on the practical implementation of pairing based cryptography. In present days, the need for identity based crypto-devices have increased. Bilinear pairing is well-suited for such applications. However it is extremely important that such devices are robust against side-channel analysis. Thus, the work also explores fault and side-channel attacks and respective countermeasures in the underlying scalar multiplication and pairing computation.

The proposed work explores efficient design techniques for pairing based cryptographic hardware. Pairing and ECSM both are formulated by a set of finite field arithmetic operations. We first propose a programmable finite field arithmetic unit. The proposed arithmetic unit can perform finite field addition, subtraction, multiplication, inversion, and division. The architectures for individual finite field arithmetic units are developed. Common and sharable circuit elements are extracted from each individual units. The common circuit elements with additional configuration logic forms the proposed programmable unit. Thereafter, we develop a hardware for scalar multiplication on respective elliptic curves using programmable finite field arithmetic unit. Side-channel attack is one of the major threats in developing cryptographic hardware. This cryptanalytic technique exploits the leakage information of the device while it executes some cryptographic algorithms. However, our proposed hardware is resistant against these type of attacks. Experimental results show that the proposed unit gives the best throughput/area ratio among contemporary designs, also considering the requirements for side-channel resistance. A programmable hardware is aimed to propose in future for computing pairings.

The high complexity of pairing implementations raises concerns regarding their reliability. Research is therefore needed to develop methodologies and techniques for designing robust system for pairing computation. It is necessary to protect the system against both accidental faults and intentional intrusions and attacks. Mainly two types of fault, namely, transient fault and permanent fault are considered in case of pairing computations. Fault injection attacks and its countermeasures on pairing based cryptography has been explicitly studied by Page and Vercauteren. We show that the countermeasures that are used for point multiplication do not work for protection of pairing algorithms against fault attacks. This work analyzes the countermeasures proposed in for pairing algorithms. It analyzes the pitfalls of the proposed countermeasures and suggests alterations to prevent fault attacks. In similar lines, the security of fault attacks against pairing computation in Edwards coordinates are also analyzed.



Satya Gautam Vadlamudi

Email: satya@cse.iitkgp.ernet.in

Satya Gautam Vadlamudi received a B.Tech.(Hons.) degree in Computer Science & Engineering along with a Minor in Mathematics and Computing from Indian Institute of Technology Kharagpur in 2008. From June 2008 till July 2009, he worked at Google India Pvt. Ltd., Bangalore, as a Software Engineer. Since August 2009, he has been a research scholar in the department of Computer Science & Engineering at IIT Kharagpur where he is also a research consultant for the General Motors Collaborative Research Lab projects on Electronics, Control and Software. His main research interests are in the areas of Computer Architectures & Systems, and Artificial Intelligence.

Supervisor: Prof. P. P. Chakrabarti

Efficient Automotive Architectures & Systems* and Efficient Heuristic Search

With the increasing number of X-by-wire systems in the automotive domain, and the share of cost of the electronic components (hardware and software) rising up to as much as 40% of the total cost of an automotive vehicle, it is important to develop efficient automotive systems that are reliable and affordable. For example, an efficient car should be the one which consumes minimum amount of fuel, giving good performance with as many systems as possible supported by wire control, providing maximum safety, requiring minimum maintenance, causing little pollution, having longer lifetime, with all the latest infotainment facilities, etc. and that has a reasonable price. On an effort to satisfy these requirements, research is being carried out in various directions in various disciplines of which electronics, control and software emerge as some of the prominent areas.

This large scale proliferation of electronic components and software as building blocks of automotive vehicles has made fault-tolerance, a first class design requirement. It is of paramount importance to develop systems which preserve functionality in spite of failures and errors in electronic, communication and processing components. This requirement stems from the fact that unlike mechanical components, which have graceful degradation in case of failures, failure of electronic components causes sharp changes in system behavior. For example, a defective mechanical steering column may cause some variations in the steering torque provided by it, but a stuck-at-fault in a processor output providing signals to a steer-by-wire system may cause dramatic increase or decrease in steering torque. Additionally, automotive systems are safety-critical systems which must conform to stringent industry safety norms for fault-tolerance. Hence, it is important to design systems which are provably resilient to faults which may emanate from various sources and causes.

Techniques have been proposed by our team (Dr. Dipankar Das, Prof. P. P. Chakrabarti et al.) to identify 'hot spots' early in the design cycle which help to re-configure the architecture so that the system satisfies various safety norms for fault-tolerance. Traditional design verification techniques for the same severely lack in coverage of various fault-occurrence possibilities due to the large simulation time involved. Our methodology efficiently characterizes

components and performs verification on that, by-passing the computationally intensive simulations, which results in increased coverage. Both static and dynamic methods for design verification are developed in this scenario. Future research is concentrated on developing self-healing architectures and fault-tolerant communication techniques for automotive vehicles amongst others.

Another area of research that we are very much involved in is Heuristic Search whose details are given below.

Heuristic search techniques are mainly studied under Artificial Intelligence where the agent's knowledge of a system is used to improve the search. These are heavily employed to solve many optimization problems that are NP-hard. For any large instance of these problems finding an optimal solution is highly expensive in terms of time and memory which has led to the development of anytime algorithms and memory-bounded algorithms respectively. For example, a 25-city Traveling Salesman Problem (TSP) has a search space of size $6.20448402 \times 10^{23}$ (greater than Avogadro number!) while a 100-city TSP has a search space of size $9.33262154 \times 10^{155}$ which explain the need for efficient heuristic search techniques that can run under limited memory conditions yet producing near-optimal results quickly.

Anytime algorithms are those which produce a sub-optimal solution quickly, and improve upon it as time progresses, giving better solutions at regular intervals. Memory-bounded algorithms are the ones which run under limited memory conditions yet producing good results. Our research is focused on developing memory-bounded anytime algorithms that are complete, those which run under limited memory conditions yet giving anytime performance, eventually terminating with optimal solution. An anytime heuristic search algorithm called Anytime Window A*(AWA*) was proposed (Dr. Sandip Aine, Prof. P. P. Chakrabarti et al.) which is complete and whose performance is comparable with the best known algorithms.

A memory-bounded version for AWA* is developed called Memory-bounded Anytime Window A*(MAWA*) which is also complete and whose performance over several benchmark TSPs is significantly better than the other well-known algorithms. Future research is concentrated on developing Memory-bounded Contract Anytime Algorithms (which produce the best possible result in the given time and memory) amongst others.

** This is part of the work being done at the General Motors Collaborative Research Lab on Electronics, Control and Software of IIT Kharagpur*



Shyamosree Pal

Email: shyamosree@cse.iitkgp.ernet.in

Shyamosree Pal received B.E. degree in Computer Science and Engineering from Birbhum Institute of Technology, Suri in 2004, and an M.E. degree in Computer Science and Engineering from Bengal Engineering and Science University, Shibpur, Howrah in 2008. Since July 2008, she has been a research scholar in the department of Computer Science and Engineering at IIT Kharagpur. Her primary research interest lies in digital geometry with applications to image analysis and computer graphics.

Supervisors: Prof. Partha Bhowmick

Estimation of Discrete Curvature Based on Number-Theoretic Characterization of Digital Curves

Opposed to the continuous nature of a curve in the real plane, a digital curve is inherently discrete in nature and requires special techniques for realization of various algorithms in the digital plane. The need to represent real objects in the digital plane has led to the emergence of the new field of digital geometry, which is different from Euclidean geometry in many remarkable aspects. As the curvature measure is an important property of an arbitrary curve, it can be used as the signature of a discrete curve for various applications in image analysis, computer graphics, and computer vision. It is found to have a relation with digital circularity, which is also a well-researched topic in the field of *digital geometry* and has several other real-world applications like circularity measure, circular arc segmentation, etc.

Estimation of discrete curvature is a challenging problem, since a mere replacement of functional derivatives by numerical differences fails to produce the desired result. Several algorithms have been proposed so far, which are mostly based on the concepts of real geometry and hence are computationally expensive. The existing measure of k -curvature, though easy to compute, is crippled with some unwanted syndromes arising out of improper consideration of chain codes. Hence, to suit the requirements, the definition of real curvature ($y'' = (1 + y'^2)^{3/2}$), as defined for a curve $y = f(x) \in \mathbb{R}^2$ is modified in our work to estimate the discrete curvature at each point on an arbitrary digital curve, C whose underlying equation is not known. Certain geometric properties of *digital straightness* have been also used to estimate the degree of bending of C in and around a point p . In effect, a smoothed version of discrete curvature is estimated, where k can be viewed as a *smoothing parameter*. The improved algorithm for estimating k -curvature is marked by its inherent simplicity and computational attractiveness, and produces the expected estimate, whether the concerned point has an extreme (high or low) curvature or the concerned segment has a constant or changing curvature (Fig. 1).

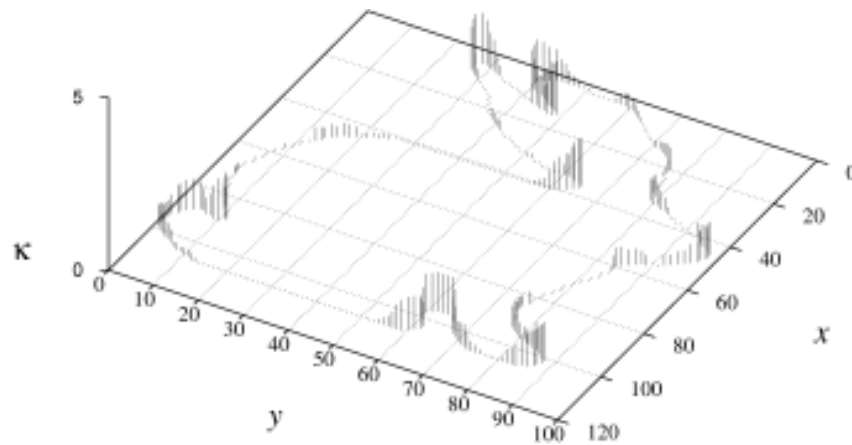


Figure 1: Curvatures estimated for a real-world digital curve representing a “rabbit” ($k = 6$) using the proposed algorithm. It may be noticed how the high curvatures (e.g., the “mouth” and the tips of two “ears”) are correctly estimated by the proposed algorithm.

Digital circularity can also be used to estimate the curvature at a point in a digital curve. This is because the curvature of a circle at every point is the reciprocal of its radius. Hence the curvature of a smooth curve is defined as the curvature of its osculating circle at each point. The same concept can be used to estimate the discrete curvature at a point in a digital curve. For this, a novel technique of deciding the digital circularity of a given digital curve segment has been developed by us. The idea is based on the correspondence of the constituent runs of digital points of a digital circle with the distribution of perfect squares (square numbers) in integer intervals. The notion of radii nesting is used to successively analyze these runs of digital points. It has been shown why and how the conflicting radii play a crucial role during the analysis and subsequently why and how the rate of convergence of the radius interval depends on the pattern of runs that constitute the digital curve segment. Two algorithms have been proposed along with their demonstrations and detailed analysis, and a simple-yet-effective solution to expedite them using an infimum circle and a supremum circle that bound the initial range of radii has been explained. That segmentation of an arbitrary digital curve segment into a sequence of circular arcs can be performed with the help of these algorithms, has been also shown in our work. Experimental results demonstrate the efficiency and elegance of the proposed technique.



Soumen Bag

Email: soumen@cse.iitkgp.ernet.in

Soumen Bag received B.E. degree in Computer Science and Engineering from REC Durgapur in 2003. From January 2004 till June 2006, he worked in Bengal College of Engineering and Technology, Durgapur, as a lecturer in the department of Computer Science and Engineering. He received M.Tech. degree in Information Technology (Department of Computer Science & Engineering) from NIT Durgapur in 2008. Since July 2008, he has been a research scholar in the department of Computer Science and Engineering at IIT Kharagpur. His research interests are in the areas of Image Processing, Document Image Analysis, and Pattern Recognition.

Supervisors: Prof. Gaurav Harit

Structural Analysis of Character Images

Designing of OCR for Indian script is a big challenge due to the presence of conjunctive characters and structural complexity of the script. For this reason, structural analysis of character images has become very promising area of research. Thinning is the most important preprocessing tool for analyzing the structural shape of character images properly and efficiently.

Thinning of shape has a wide range of applications in image processing, machine vision, and pattern recognition. But removal of spurious strokes or shape deformation in thinning is a difficult problem. In the past several decades many thinning algorithms have been developed considering all these problems. They are broadly classified into two groups: raster scan based and medial axis based. Raster scan based methods are classified into two other categories: sequential and parallel. Sequential algorithms consider one pixel at a time and visit the pixel by raster scanning or contour following. Parallel thinning algorithms are based on iterative processing and they consider a pixel for removal based on the results of previous iteration only. Many of the raster scan based character thinning methods can not preserve the local properties or features of the character images properly. The main distortion occurs at the junction regions. As a result, they give slightly shape-distorted output. Medial axis based methods generate a central or median line of pattern directly in one pass without examining all the individual pixels. They also give slight distorted result at some local regions and junction regions.

We have addressed a medial axis based thinning strategy for generating proper skeleton of character images. The algorithm is non iterative and template free. It uses shape characteristics of text to determine the areas within the image region to stop thinning partially. This approach helps to preserve the local features and true shape of the character. Additionally, it produces a set of vectorized strokes with the thin skeleton as by-product. These resultant skeletons have stronger ability to cope with shape deformation and prove more efficacious for feature extraction and classification using OCR. The advantages of our proposed technique are: (1) It attempts to minimize local distortions by making use of shape characteristics of text. (2) It provides a vectorized output of the thin skeleton

which is a collection of strokes; hence further stroke segmentation is not required. (3) Many of the spurious thinning branches which inevitably occur when applying other rasterized thinning algorithms do not occur in our proposed algorithm. (4) Our approach is most suitable for thinning text alphabets, printed or handwritten. Particularly it gives correct output even in the presence of changing width of the strokes.

At the time of evaluating the performance of thinning algorithm, we need to take care of the following criteria: (1) The algorithm must preserve the connectivity. (2) The width of resultant skeleton is one pixel. (3) The skeleton edges are not shorten. (3) The skeleton is very close to medial axis. (4) The local properties or features of the character images are preserved after the completion of thinning. (5) The thinning result is very close to the true shape of the input image. We have tested our proposed algorithm on printed English and Bengali character images. We have compared our test results with three other thinning algorithms and have observed better performance compared to all other algorithms. The proposed thinning approach has a good potential for applications to improve the performance of OCR in Indian script.

Finally, we have analyzed the structural shape of thinned images by describing the direction of enclosure of different strokes with respect to different view angle. We classify the shape of stroke into four categories: hole, convexity, concavity, and linearity. This approach helps to recognize the printed and handwritten character of different shape, font, size, and orientation.



Soumyajit Dey

Email: soumyad@cse.iitkgp.ernet.in

Soumyajit Dey received his B.E. in Electronics and Telecommunication Engg. from Jadavpur University, Kolkata in 2004 and M.S. in Computer Science and Engg. from Indian Institute of Technology, Kharagpur in 2007. Since July 2007, he has been a research scholar in the department of Computer Science & Engineering in Indian Institute of Technology Kharagpur. His research interests are in the areas of Verification and Modeling of Heterogeneous Embedded Systems.

Supervisors: Prof. Anupam Basu and Prof. Dipankar Sarkar

Formal Analysis of Heterogeneous Embedded Systems using Tagged Signal Models

The design of complex and heterogeneous embedded systems frequently requires different models of computations (MoCs) for modeling different sub-systems. In these cases, there is a need for a heterogeneous behavioral modeling technique that makes it possible to reason formally on the combination of these models, i.e., their product. The denotational framework of tagged signal model (TSM) has long been advocated as a unified modeling framework that can capture the essential features of different MoCs. A tagged signal model defines precisely processes, signals, events and provides a framework for capturing the essential properties of MoCs like discrete-event models, dataflow models, rendezvous based systems and process networks.

In the present work we embark on a two-fold objective. One is to evaluate the performance of heterogeneous designs, given an execution policy, using the model of tagged systems as an intermediate representation. The second one is to provide an algebraic characterization of the tagged systems by showing its conformance to the structure of Kleene semirings. Such a characterization helps in equational reasoning on heterogeneous specifications using the axioms of Kleene algebra.

The denotational semantics of tagged systems deals with behaviors captured by “traces” which is not a finite model of the underlying system. The theory of tag machines has recently been proposed as a finitary representation framework of the tagged systems. Tag machines are able to capture a wide range of concurrency models like asynchronous, synchronous-reactive, Time Triggered Architectures (TTA) and causality. The framework of tagged systems is based on the concept of tags. A tag structure represents an ordering relationship among events, sequences of which describe the system behaviors. Depending on the choice of the tag structure, a tag machine can capture a given concurrency model. However, such machines are not natural representations of systems specifications which is possible using formalisms like Kahn Process Network (KPN), Timed Automation (TA), Synchronous Dataflow Graph (SDFG), etc. Hence, an immediate requirement becomes formulating translation mechanisms from such specification models to tagged representations which can be composed. In this front, we have devised such a methodology for translating a given TA model to a tag machine. As a case study, we took jobshop schedules given as TAs and derived corresponding tag machines. Using the resulting compositional machine, we could derive the

asymptotic throughput of an infinite jobshop schedule which was not possible using the corresponding TA representation. Apart from TA, we have also shown the applicability of our approach to asymptotic performance evaluation of systems modeled using SDFGs. Further, we have verified the applicability of our approach in case of heterogeneous systems by evaluating the performance of a system comprising of dataflow and discrete-event blocks.

In order to provide an algebraic characterization of tagged systems, the second part of the work builds on domain theory, developed for the denotational semantics of programming languages. Actor oriented formalisms established in the last two decades are based on such concepts. Similarly, in the tagged signal model, we have the concepts of actors mapping a set of input signals to a set of output signals. In the original model of tagged signals (popularly known as LSV named after Lee and Vincentelli), the set of all such tagged signals has been shown to form a complete partial order (CPO). We show that a similar result can be derived in the context of the more succinct version of the model which is proposed in [1]. Our actors support bounded, value-based, biased non-determinism. We further prove that the set of all such possible TSM actors is closed under the axioms of Kleene algebra (KA) and its extensions like Kleene algebra with Test (KAT) and Kleene Algebra with Domain (KAD) [2]. The result has important consequences. For a given a heterogeneous system specification, we can transform it into a corresponding TSM actor based representation which can be encoded as an algebraic expression of KA/KAT/KAD. Using the axioms of these algebras, we can perform property verification as well as functional equivalence checking of such complex systems. As a case study focussed on equivalence checking of heterogeneous embedded systems, we construct two different implementations of a Reflex Game modeled using the Synchronous Reactive (SR) MoC, derive the corresponding KAT based encodings and prove their equivalence.

As part of future work, we intend to model control intensive heterogeneous systems using our formalism and perform formal property verification of such systems by using automated theorem provers guided by the algebraic axioms of KA/KAT/KAD.

References

1. **A. Benveniste, B. Caillaud, L. Carloni, P. Caspi, and A. Sangiovanni-Vincentelli.** Composing heterogeneous reactive systems. ACM TECS, 7(4), 2008.
2. **Dexter Kozen.** Kleene algebra with tests. ACM Trans. Program. Lang. Syst., 19(3):427–443, 1997.



Sourav Das

Sourav Das received a B.Tech. degree in Instrumentation Engineering from Indian Institute of Technology, Kharagpur in 1999, and an M.S. degree in Telecommunication and Software Engineering from BITS, Pilani in 2004 (Off Campus). From July 1999 till date, he worked in Wipro Technologies, Lucent Technologies, Huawei Technologies and Alcatel-Lucent in various roles such as Software Developer, Lead Engineer, Project Manager and Solution Architect. Since January 2007, he has been a sponsored research scholar in the department of Computer Science & Engineering in IIT Kharagpur. His research interests are in the areas of Cryptography and Network Security.

Supervisor: Prof. Dipanwita Roychowdhury

Studies on Cellular Automata Based Stream Ciphers

Stream Ciphers are an important class of symmetric key cryptography. The objective of stream ciphers is to provide ultra-fast encryption that can be used in communication channels. However, most of the classic stream ciphers have been broken due to their simplistic design to achieve higher speed. In search of a good stream ciphers ESTREAM project has been launched.

Cellular Automata, due to their speed of execution and statistical randomness, can be a very good building block for stream ciphers. But, unfortunately, the Cellular Automata, unlike LFSR, have not gained much attention in design of stream ciphers. However, it has been shown that Cellular Automata can provide better resistance against correlation and leakage of information than LFSR.

Grain is one such stream cipher that uses LFSR and is among the finalist candidates of ESTREAM. There were a few attacks reported on Grain. We have analyzed that all the reported attacks on Grain are due to the weakness of LFSR. We have shown that, if we replace the LFSR block in Grain with a maximum length Cellular Automata, then all the reported attacks can be prevented.

Hiji-bij-bij is another stream cipher that was proposed in 2003 and uses Cellular Automata. However, Hiji-bij-bij was broken within a short time and five attacks have been reported on Hiji-bij-bij. We have analyzed the Hiji-bij-bij stream cipher and each of the five attacks reported on it. We have found that none of the weaknesses of Hiji-bij-bij is because of Cellular Automata. The weaknesses are due to the improper design of its non-linear block, which does not use Cellular Automata. We have redesigned the non-linear block of Hiji-bij-bij with the Cellular Automata based larger S-boxes developed by us. We showed that the modified Hiji-bij-bij with the new scheme is able to prevent all the five reported attacks.

One of the disadvantages of Cellular Automata is that it is completely linear. However, Non-Linearity is one of the most important criteria of any cryptographic algorithm. It is possible to generate non-linear CA using non-linear

rules, but the problem is such non-linear cellular automata have very short period length and it is impossible to utilize the whole state space. We have devised a method to generate non-linear maximum length CA as a part of our research.

The stream ciphers require large, efficient S-boxes since their algorithms are simple and prone to attacks. However, there is no generic method that can generate large scalable S-boxes. Using the non-linear CA as a non-linear primitive, we have given a schema to generate large, scalable and parameterized S-boxes that are suitable for stream ciphers. These S-boxes are easy to implement both in hardware and software and run efficiently. We have analyzed the security properties of the S-boxes in terms of Non-linearity, Algebraic Degree, number of terms and degree distribution in the Algebraic Normal Form of the co-ordinate functions, Strict Avalanche Criterion, Linear Cryptanalysis and Differential Cryptanalysis. We found that the S-boxes generated using our scheme have excellent security properties.

Currently, we are working on creating a stream cipher completely based on Cellular Automata. We have devised a scheme for such a stream cipher. We have analyzed the security properties for a small scale version of the stream cipher and the results are promising. We are evaluating the stream cipher against all the known attacks to make it acceptable for the community.



Sourav Kumar Dandapat

Sourav Kumar Dandapat received a B.E. degree in Computer Science from Jadavpur University in 2002, and an M.Tech. degree in Computer Science from IIT Kharagpur in 2005. From July 2005 till November 2007, he worked in IBM ISL, Bangalore, as a System Software Engineer. From December 2007 till February 2009, he worked in Magma Design Automation, Bangalore, as an associate member of technical staff. Since July 2009, he has been a research scholar at the department of Computer Science and Engineering in IIT Kharagpur. His research interests are in the areas of Wireless Internet.

Supervisor: Prof. Niloy Ganguly

Fair Load Balancing in Wireless Mobile Environment Using Max-Flow

Wireless clients must associate to a specific Access Point (AP) to communicate over the Internet. Current association methods are based on maximum Received Signal Strength Index (RSSI) implying that a client associates to the strongest AP around it. This is a simple scheme that has performed well in purely distributed settings. Modern wireless networks, however, are increasingly being connected by a wired backbone. The backbone allows for out-of-band communication among APs, opening up opportunities for improved protocol design. We want to take advantage of this opportunity through a coordinated client association scheme where APs consider a global view of the network, and decide on the optimal client-AP association. We show that such an association outperforms RSSI based schemes in several scenarios, while remaining practical and scalable for wide-scale deployment. Although an early work in this direction, our basic analytical framework (based on a *max-flow* formulation) can be extended to sophisticated channel and traffic models. Our future work is focused towards designing and evaluating these extensions.

So here we are going to propose a load balancing algorithm for wireless mobile environment to increase the fairness and overall throughput of the network using maximum flow algorithm.

We consider an area (say an institute campus), which is entirely covered by wireless Internet. Every location in that campus is either covered by a single AP or multiple APs. We denote an area as *exclusive zone* if that area is in the vicinity of a single AP. Similarly we denote an area as *shared zone* if the area is covered by multiple APs.

An AP broadcasts beacon packets every *beacon interval period*. In response, mobile devices (interested for connection) may either send a *probe request* or may ignore them.

We assume for a zone z_i total number of device is noted as t_i , and number of admitted device using an association control algorithm is s_i . We also assume there are m APs and n zones.

Most important requirement in any load balancing protocol is fairness in allocation. Fairness is measured as the standard deviation from mean zone satisfaction (MZS), where formal definition of MZS and Fairness Deviation (FD) are

$$MZS = \frac{\sum_{i=1}^n \frac{s_i \times 100}{t_i}}{n} \quad FD = \sqrt{\frac{\sum_{i=1}^n (MZS - s_i \times 100 / t_i)^2}{n}}$$

The proposed algorithm also aims to maximize the Percentage of Connection Admitted (PCA) of the network. A device needs a unit of bandwidth to get admitted. PCA is defined formally as

$$PCA = \frac{\sum_{i=1}^n s_i \times 100}{\sum_{i=1}^n t_i}$$

With our scheme, a client temporarily associates to the AP using RSSI-based association. As a part of the association process, the client also conveys the IDs of all APs that it has recently overheard. The AP shares this information with other APs (over the wired backbone), thereby learning about the population of clients in different zones around each AP. This per-zone population is used to facilitate better association, as described in our protocol next.

Now, the key idea is to periodically execute an optimal association algorithm (every *mapping interval*) and re-assign the clients to APs for improved performance. In effect, each AP reallocates its own resources among its influencing zones, so that desired metrics are optimized. This reallocation is a function of the current traffic distribution, i.e., the load on the network. We use a 4-stage max-flow and this process is termed as Fair Bandwidth Allocation (FBA) algorithm. After the re-allocation is complete, every AP enforces this optimal allocation through a minimal number of re-associations. In between two executions of FBA devices associate using RSSI.

To evaluate the performance of FBA we compare FBA with one RSSI based algorithm and one LLF based algorithm. In LLF based association device associates with least loaded AP first. Comparison with benchmark algorithms shows that our algorithm performs better than traditional algorithms.

Here we propose to improve client association techniques in wireless networks by exploiting the wired backbone among WiFi APs. The key idea is to share local information from multiple APs, model it as a max-flow problem, and derive the optimal client-to-AP assignment. Simulation results demonstrate that such a technique can improve over purely distributed association schemes, resulting in higher fairness, better load balancing properties, and even some robustness to client mobility. We have made several assumptions that may not be precise in real conditions. Our future work is focused towards relaxing these assumptions and augmenting the system with sophisticated channel and traffic models.



Srobona Mitra

Email: srobona@cse.iitkgp.ernet.in

Srobona Mitra received a B.E. degree in Computer Science and Engineering from Jadavpur University, Kolkata in 2004 and an M.Tech. degree in Computer Science and Engineering from Indian Institute of Technology Kharagpur, Kharagpur in 2006. From June 2006 till March 2007, she worked in IXIA Technologies Pvt. Ltd., Kolkata as a Software Development Engineer. Since April 2007, she has been a research scholar in the Department of Computer Science and Engineering in Indian Institute of Technology Kharagpur. Her research interests are in the areas of VLSI design and Verification.

Supervisors: Prof. Pallab Dasgupta and Prof. P. P. Chakrabarti

Formal Methods for Incremental Verification

With increasing complexity in the field of VLSI, design validation takes up more than 70% of the design cycle time of most chips. Traditionally, there are three main types of design verification paradigms, namely: (a) Simulation-based Verification, (b) Formal Property Verification (FPV) and (c) Sequential Equivalence Checking (SEC), each of which has its own applicability in certain domains.

Simulation of a design results in traces which are sequences of valuations to different signals of the design in each simulation cycle. On the other hand, the results of FPV are proven properties on hierarchically specified designs and the results of Sequential Equivalence Checking are established invariants. These results carry significant amount of information on the structure and behavior of the designs under verification (DUV). Typically, these results are not reused in validation. When a new verification problem over the same DUV comes in, such as proving new properties on the existing DUV, or proving existing properties on a partially modified implementation of the same DUV, or both, verification is started from scratch, without using the existing results. In this work, our objective is to devise validation methods that reuse prior verification results to solve new verification problems on the same DUV incrementally and efficiently.

A large scale industrial circuit normally consists of a large number of blocks, each of which may consist of sub-blocks and so on. Each such component block has local formal properties and simulation traces. The global design which encompasses all these blocks and sub-blocks also has its own global architectural properties and some simulation traces which may have been run globally on the integrated design. Our objective is to build over this integrated design a framework for incremental verification. We consider the following problems in this research:

1. Verification by parts: Reusing Component Invariant Checking Results: In this problem, we consider reuse of previous verification results in the domain of SEC. The verification problem in this domain reduces to checking that no state is reachable where an output of the state machine model differs with that of the design. The proof establishes the invariant that the outputs always match. Given a set of proven (local) invariants over the components

of an integrated design, and a (global) invariant which we want to prove over the integrated design, our challenge is to devise a formal method to use the known local invariants to cut down the search space for the global invariant. Our work addresses this problem by proving invariants by BDD-based backward reachability analysis. Experimental results show that our method achieves an average of 39% decrease in transition relation BDD size and also significantly decreases the number of cycles required to prove a global invariant on large-sized designs.

2. Incremental approach to verification for a change in specification. In this problem we consider the scenario where a new architectural property is introduced into the design hierarchy. Our objective is to verify whether this new property holds on the design reusing the existing results, that is, the proved properties and the traces.
3. Methodology for environment model generation for an arbitrary cone of logic. In order to formally verify a new property at any intermediate level of hierarchy of a design, which has been introduced due to a failure trace being encountered, our objective is to identify a cone of influence of the property appropriately, so that we can verify the property on this cone only instead of the entire design. So the question arises as to what sequences of valuations should be assumed for the interface of this cone with the rest of the design. In order to avoid spurious failures of the property, our objective is to generate an environment model for this cone for verification. Verification of the property on this cone with this environment model after bug fix should ensure that this bug and all related bugs of the same family have been eliminated from the design.
4. Incremental verification for a change in both the specification and the design (bug fix). Our objective is to propagate a new property, conceived due to a failure trace, added at some level of hierarchy of the design, both upwards and downwards along the design hierarchy, so that the entire framework gets updated with the implications of the new property. When a failure trace is obtained at some intermediate level module, then a new property is added to its local property suite and a design change is done in the module to fix the bug. The joint effect of these two changes is that many of the lower level properties of modules that constitute this affected module can conflict with the new property. Also many of the existing local properties of these modules, which had been proved locally on the modules, can get violated now because of the design fix. Now our objective is to stabilize the framework so that all the conflicting lower level properties are eliminated from the system, keeping the ones which are not conflicting, so that they are proved on the lower level modules.

This research was partially supported by the IBM Faculty Award of Prof Pallab Dasgupta and is currently being partially supported by a research grant from Intel.



Subhankar Mukherjee

Subhankar Mukherjee received an M.Tech. and a B.Tech. degree in Electrical Engineering from Indian Institute of Technology Kharagpur, under the dual degree program, in 2008. Since July 2008, he has been a research scholar in the Department of Computer Science and Engineering in Indian Institute of Technology Kharagpur. His research interests are in the areas of VLSI Modeling and Verification with a focus on Analog Assertion-Based Verification.

Supervisors: Prof. Pallab Dasgupta and Prof. Siddhartha Mukhopadhyay (EE)

Design Intent Verification of Mixed-Signal Systems

Simulation has been the primary technology for validating the integration of a heterogeneous collection of analog/digital design components (IPs) into an integrated circuit. Given the spiraling complexity of present day mixed-signal system-on-chip designs and the lack of proportionate growth in the speed of mixed-signal simulation, it is becoming increasingly infeasible to achieve adequate coverage by simulation. One way to overcome the difficulties in integrating a system is to raise the level of abstraction at which the integration is performed, primarily to facilitate the verification. In other words, we substitute the components with meta-models that capture the intent of the components, verify the integrated design functionality by integrating the meta-models, and finally verify that each component is acceptable with respect to its meta-model. In the digital domain, several formalisms have emerged for modeling the design intent. This includes abstract state machines, temporal properties and Boolean formulas. In the mixed-signal domain, similar abstractions include hybrid automata, and time and frequency domain properties. Formal design intent modeling of analog and mixed-signal designs has so far remained a myth, but it is becoming an increasingly important requirement. The focus of this research is to explore this direction. The industry trend appears to be moving towards designs that integrate large digital circuits with multiple analog/RF interfaces. In the verification of these large integrated circuits, the number of nets that need to be monitored has been growing rapidly. Consequently the mixed-signal design community has been feeling the need for mixed-signal assertions that can automatically monitor conformance with expected time-domain behavior and can help in debugging deviations from the design intent. The main challenges in providing this support are (a) developing mixed-signal assertion languages, and (b) developing support for assertion verification during mixed-signal simulation. Mixed-signal assertions is an important item in this research agenda. However, we also need to look beyond assertions for capturing, formalizing and composing the design intent of mixed-signal components.

The key objectives towards formalizing a Mixed-Signal Assertion Language are as follows:

1. To define the formal semantics of sampling the analog signals, and the possibility of handling multiple sampling clocks. Since our intent is to capture behaviors in the continuous domain, it is imperative that we should not miss

the behaviors relevant to a particular property by using coarse grained sampling. On the other hand, a fine grained sampling approach may severely affect the simulation speed.

2. To develop a prototype that can be integrated with existing multi-mode mixed-signal simulation tools through a set of APIs.
 - (a) **Developing Methods for Mixed-Signal Coverage Analysis:** One of the major benefits of assertions in the digital domain has been towards improving the notion of functional coverage. Assertions capture different corner case behaviors, and coverage monitors can report whether the scenarios relevant to a given assertion were visited during simulation. For more details on this approach, one may refer to the notion of cover properties in the System Verilog Assertion (SVA) standards. In analog domain, behavioral coverage is largely done manually through visual inspection of the test cases. This research aims to bring some formal rigor to coverage analysis in the mixed-signal domain. The key objectives of this part are to develop coverage monitors for mixed-signal assertions.
 - (b) **Formal Design Intent Modeling of Mixed-Signal Designs:** While formal properties are useful to formally express some of the key behavioral requirements of a design, it is also widely accepted that properties are too restrictive to capture the overall design intent of a circuit. Consequently, assertion IPs in the digital domain use formalisms such as auxiliary state machines to come up with a more expressive and more readable version of the design intent. In the analog and mixed-signal domain, formalisms such as hybrid automata have been studied for modeling the design intent of a circuit. Unfortunately the gain in expressibility comes with a proportionate increase in the complexity of analysis. For example, reachability in hybrid automata has been shown to be undecidable except in very special cases. Therefore, there is a need to explore formalisms for capturing the design intent of mixed-signal circuits that are expressible, and more amenable for mathematical analysis. The last part of this research aims to explore such formalisms in the mixed-signal domain.

This research is being supported by the Semiconductor Research Corporation (SRC) under the GRC contract no: 2008-TJ-1835. The research is being mentored by Freescale Semiconductor.



Sudip Roy

Email: sudipr@cse.iitkgp.ernet.in

Sudip Roy received B.Sc. with Physics honors and B.Tech. in Computer Science and Engineering both from University of Calcutta, Kolkata in 2001 and 2004, respectively, and M.S. in Computer Science & Engineering from IIT Kharagpur in June 2009. From November 2004 till July 2005, he worked with Machine Intelligence Unit of Indian Statistical Institute, Kolkata as a Project-Linked-Personnel in a research project. He worked with the Department of Computer Science and Engineering in IIT Kharagpur as a Junior Project Assistant in a research project titled “Low Power VLSI Circuits and Systems”, sponsored by INTEL Corporation, USA from August 2005 till June 2008 and pursued his M.S. (by research) degree there. From July 2008 till September 2009 he worked with the Advanced Computing and Microelectronics Unit of Indian Statistical Institute, Kolkata as a Project-Linked-Personnel in the research project titled “Floorplan Optimization for Nano-Biochips”. Since October 2009, he has been a research scholar in the department of Computer Science and Engineering at IIT Kharagpur. His areas of interest are in computer-aided design (CAD) algorithms for design and testing of digital microfluidic biochips and VLSI chips.

Supervisors: Prof. P. P. Chakrabarti and Prof. Bhargab B. Bhattacharya (ISI Kolkata)

CAD Algorithms and Techniques for Digital Microfluidic Biochips

Microfluidic-based biochips are soon revolutionizing clinical diagnostics and other biochemical laboratory procedures to meet the challenges of healthcare cost for cardiovascular diseases, cancer, diabetes, and global HIV crisis, etc. [1, 2]. A marriage of microelectronics and in-vitro diagnostics areas leads to a new field of “Lab-On-a-Chip (LOC)” or nano-biochips. Research in this new discipline needs the integration of many disciplines such as microelectronics, (bio)chemistry, in-vitro diagnostics, computer-aided design (CAD) and optimization, microchip fabrication technology, etc. [1, 5]. Typically, an LOC implements one or more biochemical laboratory protocols or assays on a single chip that is a few square centimeters in size. The emerging application areas include among others, clinical diagnostics, especially the immediate point-of-care diagnosis of diseases, enzymatic analysis (e.g., glucose and lactate assays), DNA analysis (e.g., PCR and nucleic acid sequence analysis), proteomic analysis involving proteins and peptides, immunoassay, and environmental toxicity monitoring [1, 2].

One category of microfluidic chips are continuous-flow microfluidic chips, where continuous liquid flow through microfabricated channels is manipulated with the help of micropumps, microvalves, etc. A more versatile category of biochips are digital microfluidic (DMF) biochips, where discrete and independently controllable droplets of micro/nano/pico litre volume of the biosample and reagent fluids are manipulated on a substrate of two dimensional array of electrodes using electrical actuation (a principle called electrowetting-on-dielectric or EWOD) [1, 5, 6]. Compared to traditional bench-top procedures, DMF chip technology offers the advantages of low sample and reagent consumption, less likelihood of error due to minimal human intervention, high throughput and high sensitivity, portability, increased automation, low power consumption, low cost and reliability. As each droplet (or group of

droplets) can be controlled individually, these types of biochips also have dynamic reconfigurability and architectural scalability. In general, a DMF biochip functionality includes the following operations: measuring and dispensing accurate amounts of sample/reagent fluid, transporting fluid droplets to appropriate locations, mixing of droplets, splitting of larger droplets into smaller ones, detection and analysis sample; and it can integrate multiple bioprotocol operations on a single chip [1, 5].

To build a biochip efficiently, several associated combinatorial optimization and CAD problems need to be solved. Recently, many CAD algorithms and techniques are being developed for both design and testing of DMF biochips [4, 8, 9]. Several combinatorial and geometric optimization problems arise during the computer-aided design and testing of such chips [3]. We will provide an overview of some of the recent research works in this field. At first, the descriptions and architectures of DMF biochips will be discussed. Then some CAD optimization problems involved in the design and testing of such biochips will be presented. Finally, a recent work will be discussed that is on the development of abstraction layers for microfluidic biochips, i.e., a portable programming language (BioStream) for describing biology protocols and a stable interface or instruction set architecture (Fluidic ISA) for microfluidic chip designers [7].

Since off-chip sample processing and sample preparation pose a significant hindrance to the overall biochemical assay time, for fast and high throughput applications, sample preprocessing steps should also be automated on-chip i.e., integrated and self-contained on the biochip itself. Currently, we are working on the CAD problems and issues involved in the automation of on-chip sample/reagent preparation and preprocessing steps of a bioassay, such as automated dilution and ratioed mixing of sample/reagent fluids. As a future plan, we intend to develop techniques for the design automation of DMF biochips and to solve several other combinatorial optimization problems in the design and testing of such chips.

References

1. **K. Chakrabarty**, and **F. Su**, "Digital Microfluidic Biochips: Synthesis, Testing and Reconfiguration Techniques". CRC Press, 2007.
2. **R. B. Fair** et al., "Chemical and Biological Applications of Digital-Microfluidic Devices", IEEE Design & Test of Computers, Vol. 24, No. 1, pages 10-24, 2007.
3. **K. Chakrabarty**, and **J. Zeng**, "Design Automation for Microfluidic-Based Biochips", ACM JETC, Vol. 1, No. 3, pages 186-223, October 2005.
4. **T. Xu** et al., "Design and Optimization of a Digital Microfluidic Biochip for Protein Crystallization", in Proc. of ICCAD, USA, pages 297-301, 2008.
5. **R. B. Fair**, "Digital Microfluidics: Is a True Lab-on-a-Chip Possible?", Microfluidics and Nanofluidics, Vol. 3, pages 245-281, March 2007.
6. **M. Abdelgawad**, and **A. R. Wheeler**, "The Digital Revolution: A New Paradigm for Microfluidics", Advanced Materials, Vol. 21, pages 920-925, 2009.
7. **W. Thies** et al., "Abstraction Layers for Scalable Microfluidic Biocomputing", Natural Computing, Vol. 7, No. 2, pages 255-275, May 2008.
8. **J. Ding** et al., "Scheduling of Microfluidic Operations for Reconfigurable Two-dimensional Electrowetting Arrays", IEEE TCAD, Vol. 20, No. 12, pages 1463-1468, December 2001.
9. **M. Cho** and **D. Pan**, "A High-Performance Droplet Routing Algorithm for Digital Microfluidic Biochips", IEEE TCAD, Vol. 27, No. 10, pages 1714-1724, October 2008.



**OUR MENTORS:
FACULTY OF THE DEPARTMENT**



Indranil Sengupta

Research Interests: *Cryptography and Network Security, VLSI Design and Testing, Mobile Computing*

Dr. Indranil Sengupta obtained his B.Tech., M.Tech. and Ph.D. degrees in Computer Science and Engineering from the University of Calcutta. He joined Indian Institute of Technology, Kharagpur, as a Lecturer in 1988, in the Department of Computer Science and Engineering, where he is presently a Professor and the Head of the Department. He is also heading the School of Information Technology of the Institute. A Centre of Excellence in Information Assurance has been set up at IIT Kharagpur under his leadership, where a number of security related projects are presently being executed. He has over 24 years of teaching and research experience, and over 100 publications in international journals and conferences. His research interests include cryptography and network security, VLSI design and testing, and mobile computing.



Anupam Basu

Research Interests: *Cognitive Science and Language Processing with particular focus on Intelligent Interface Design and Human Computer Interaction*

Prof. Anupam Basu is a Professor at the Dept. of Computer Science & Engineering, IIT Kharagpur, India. He has been in the faculty since 1984. His research interests include Intelligent Systems, Embedded Systems and Language Processing. His research has been directed to develop a number of cost effective Assistive Systems for the physically challenged as well as for development educational systems for the rural children. In all these applications, he has synthesized his research to lead to products, which are presently in use in several village knowledge centers as well as in several organizations for the physically challenged. He is considered to be a pioneer in Assistive Technology research in India.

Presently, he is also serving as the Director of the Society for Natural Language Technology Research, an R& D institute aimed at carrying out language localization research and development.

Prof. Basu had taught at the University of Guelph, Canada, University of California, Irvine and at the Dortmund University, Germany. He is an Alexander von Humboldt Fellow and a Fellow of the Indian National Academy of Engineering.

He has won several awards and honors for his research contributions. These include the National Award for the Best Technology Innovation for the Physically Disabled (2007), the Da Vinci Award 2004, Outstanding Young Person Award 1996.



Partha Bhowmick

Research Interests: *Digital Geometry, Shape Analysis, Computer Graphics*

Partha Bhowmick graduated from the Indian Institute of Technology, Kharagpur, India, and received his master's and PhD degrees from the Indian Statistical Institute, Kolkata, India. He is currently working as an Assistant Professor in the department of Computer Science and Technology, Indian Institute of Technology, Kharagpur, India. His primary research interest is digital geometry, pertaining to algorithms in the digital paradigm and involving potential applications in computer graphics, low-level image processing, approximate pattern matching, shape analysis, GIS, and biometrics. He has published over 45 research papers in international journals, edited volumes, and refereed conference proceedings, and holds three US~patents.



Goutam Biswas

Research Interests: *Theoretical Computer Science, Compilers.*



Partha Pratim Chakrabarti

Research Interests: *Artificial Intelligence, Algorithms for Design Automation in VLSI and Embedded Systems*

Partha P Chakrabarti is a Professor in the Department of Computer Science and Engineering at Indian Institute of Technology, Kharagpur. Currently he is also holding the post of Dean SRIC (Sponsored Research and Industrial Consultancy) and Head of the Advanced Technology Development Centre (ATDC) at IIT Kharagpur. He received the Bachelor's degree in Computer Science from IIT Kharagpur, India, in 1985. His Ph.D., in Computer Science & Engineering from IIT Kharagpur. His specific interests include Heuristic and Exploratory Search Techniques, Automated Problem Solving and Reasoning, Algorithms for Synthesis and Verification of VLSI Systems, Scheduling, Verification and Fault Tolerance Analysis of Multi-Processor Embedded Systems, etc. He has over 200 publications, and has supervised around 16 Ph.Ds. He is the principal investigator of several research projects, and is a consultant to industry and government. He helped found the Advanced VLSI Design Laboratory and the General-Motors-IIT-Kharagpur Collaborative Research Laboratory on ECS at IIT Kharagpur. As Dean SRIC, he has helped grow the sponsored research at IIT Kharagpur multiple-fold including setting up of several Advanced Research Centres of Excellence and the Entrepreneurship Programme. He is a Fellow of Indian National Science Academy, Indian Academy of Science, Indian National Academy of Engineering and The West Bengal Academy of Science & Technology. He is the recipient of several awards, including the President of India Gold Medal, Shanti Swarup Bhatnagar Award, Swarnajayanti Fellowship, INSA Young Scientist Award, Indian National Academy of Engineering (INAE) Young Engineer Award, Anil Kumar Bose Award from INSA, Best Paper Awards in International Conference on VLSI Design and National Scholarship.



Abhijit Das

Research Interests: *Arithmetic and Algebraic computations with Specific Applications to Cryptology*

Abhijit Das is Assistant Professor, Department of Computer Science and Engineering, Indian Institute of Technology Kharagpur. He has held academic positions at the Indian Institute of Technology Kanpur and Ruhr-Universität Bochum, Germany. His research interests include arithmetic and algebraic computations with specific applications to cryptology.



Pallab Dasgupta

Research Interests: *Formal Verification, Artificial Intelligence and VLSI.*

Dr. Pallab Dasgupta did his B.Tech, M.Tech and PhD in Computer Science from the Indian Institute of Technology Kharagpur. He is currently a Professor at the Dept. of Computer Sc. & Engg, I.I.T. Kharagpur. His research interests include Formal Verification, Artificial Intelligence and VLSI. He has over 100 research papers and 2 books in these areas. He currently leads the Formal Verification group at the CSE Dept., IIT Kharagpur (<http://www.facweb.iitkgp.ernet.in/~pallab/forverif.html>) which has been developing validation technology for several companies, including Intel, Synopsys, General Motors, SRC and National Semiconductors. Since Oct 2007, he is also the Professor-in-charge of the Advanced VLSI Design Lab, IIT Kharagpur. Dr Dasgupta has been a recipient of the Young Scientist awards from the Indian National Science Academy, Indian National Academy of Engineering, and the Indian Academy of Science. He is a senior member of IEEE.



Partha Sarathi Dey

Research Interests: *Digital Logic Design, Data Structures, Computer Organization and Architecture*

M.Tech.(IIT Kharagpur)

P S Dey joined the Institute in 1985



Niloy Ganguly

Research Interests: *Peer-to-peer Networks, Complex Network Theory, Social Networks Modeling*

Niloy Ganguly is an associate professor in the department of computer science and engineering, Indian Institute of Technology Kharagpur. He has received his PhD from Bengal Engineering and Science University, Calcutta, India and his Bachelors in Computer Science and Engineering from IIT Kharagpur. He has been a post doctoral fellow in Technical University of Dresden, Germany where he has worked in the EU-funded project Biology-Inspired techniques for Self-Organization in dynamic Networks (BISON). He presently focuses on dynamic and self organizing networks especially peer-to-peer networks, web-social networks, delay tolerant network etc. In peer-to-peer networks he has worked on optimizing various services like search, topology management and applications like IP telephony, publish-subscribe system etc. In social networks, he has worked on designing recommendation system based on community structures on various web-social networks like Twitter and Delicious. He has also simultaneously worked on various theoretical issues related to dynamical large networks often termed as complex networks. In this line he has been instrumental in organizing the workshop series Dynamics on and of Complex Networks in European Conference on Complex Systems. He has published around 70 papers in international conferences and journals. He has also edited a book on Complex Networks published by Birkhauser, Boston.



Sujoy Ghose

Research Interests: *Design of Algorithms, Artificial Intelligence, and Computer Networks*

Sujoy Ghose received the B.Tech. degree in Electronics and Electrical Communication Engineering from the Indian Institute of Technology, Kharagpur, in 1976, the M.S. degree from Rutgers University, Piscataway, NJ, and the Ph.D. degree in computer science and engineering from the Indian Institute of Technology. He is currently a Professor in the Department of Computer Science and Engineering, Indian Institute of Technology. His research interests include design of algorithms, artificial intelligence, and computer networks.



Arobinda Gupta

Research Interests: Distributed Systems, Networks

Arobinda Gupta received his Ph.D. in Computer Science from the University of Iowa, Iowa City, in 1997, an M.S. in Computer Science from the University of Alabama in 1992, and an M.E. and a B.E. in Electronics and Telecommunication Engineering from Jadavpur University, Kolkata, India in 1990 and 1987 respectively. From February 1997 to September 1999, he was with the Windows 2000 Distributed Infrastructure group in Microsoft Corp., Redmond, Washington, USA. Since Oct. 1999, he is a faculty in Indian Institute of Technology Kharagpur, where he is currently an Associate Professor in the Department of Computer Science & Engineering and School of IT. His current research interests are broadly in the areas of distributed systems and networks, and more specifically in adaptive distributed systems, vehicular ad hoc networks, and delay tolerant networks.



Gaurav Harit

Research Interests:

- *Video Analysis: Event Detection, Audio analysis, Semantic Modeling, Video Ontology*
- *Computer Vision: Shape Modeling, Texture Analysis*
- *Document Image Analysis: Segmentation, Layout Analysis, Word Image-based Indexing*
- *Pattern Classification: Optical Character Recognition for Indian Scripts*

Gaurav Harit received his B.E. in Electrical Engineering from MBM Engineering college, JNV University Jodhpur in 1999. He received the Master's Degree in Computer Technology from Electrical Engineering Department, Indian Institute of Technology (IIT) Delhi in 2000. He received his PhD degree from IIT Delhi in 2007. Since 2008 he has been working as an Assistant Professor in the department of Computer Science and Engineering at IIT Kharagpur. His research interests include Computer Vision, Multimedia Systems, and Artificial Intelligence.



Rajeev Kumar

Research Interests: *Programming Languages & Software Engineering, Embedded & Multimedia System, Evolutionary Computing.*

Rajeev Kumar received his Ph.D. from University of Sheffield and M.Tech. from University of Roorkee (now, IIT Roorkee) both in computer science and engineering. Currently, he is a professor of computer science and engineering at IIT Kharagpur. Prior to joining IIT, he was with the Birla Institute of Technology & Science (BITS), Pilani and the Defense Research and Development Organization (DRDO). His research interests include programming languages & software engineering, embedded & multimedia system, and evolutionary computing for combinatorial optimization. He has supervised 8 Ph.Ds and published over 150 research articles. He is a senior member of ACM and IEEE, and a fellow of IETE.



Arun Kumar Majumdar

Research Interests: *Data and Knowledge-based Systems, Multimedia Systems, Medical Informatics, VLSI Design Automation.*

A. K. Majumdar obtained B. Tech, M. Tech and Ph. D degrees in Applied Physics from the University of Calcutta in 1967, 1968 and 1973, respectively. He also obtained a Ph. D. degree in Electrical Engineering from the University of Florida, Gainesville, U. S. A., in 1976. Since 1980, he is associated with the Indian Institute of Technology, Kharagpur, first as an Assistant Professor in the Electronics and Electrical Communication Engineering Department and then from 1984 as a Professor in the Computer Science and Engineering Department. With leave from IIT, Kharagpur, he served as a Visiting Professor in the University of Guelph, Ontario, Canada in 1986-87, and in the George Mason University, Fairfax, Virginia, USA, in the summer of 1999. Earlier, he worked in the Indian Statistical Institute, Calcutta, and Jawaharlal Nehru University, New Delhi, as a faculty member. He is currently the Deputy Director, IIT Kharagpur. He has also served as Head, School of Medical Science & Technology, IIT Kharagpur, from 2005 to 2006, Dean (Faculty and Planning), IIT Kharagpur from March 2002 to 2005, Head of the Computer Science and Engineering Department, IIT Kharagpur from 1992 to 1995 and again from 1998 to May 2001 and Head of Computer and Informatics Center, IIT Kharagpur: from 1998 to 2002.



Rajib Mall

Research Interests: Program Analysis and Testing

Rajib Mall has been with the Computer Science and Engineering at IIT, Kharagpur since in 1994. Prior to joining IIT, Kharagpur, he worked with Motorola India for about three years. Dr. Mall completed all his professional education: Ph.D., Master's, and Bachelor's degrees from the Indian Institute of Science, Bangalore. He has guided 12 Ph.D. dissertations and has authored two books. He has published more than 150 research papers in International refereed conferences and Journals. Dr. Mall works mostly in the area of program analysis and testing.



Chittaranjan Mandal

Research Interests: Formal Modelling and Verification, High-level Design, Network and Web Technologies

Chittaranjan Mandal received his Ph.D. degree from IIT, Kharagpur, India, in 1997. He is currently an Associate Professor with the Department of Computer Science and Engineering and also the School of Information Technology, IIT, Kharagpur. Earlier he served as a Reader with Jadavpur University. His research interests include formal modelling and verification, high-level design and network and web technologies. He has about seventy publications and he also serves as a reviewer for several journals and conferences. Prof. Mandal has been an Industrial Fellow of Kingston University, UK, since 2000. He was also a recipient of a Royal Society Fellowship for conducting collaborative research. He has handled sponsored projects from government agencies such as DIT, DST and MHRD and also from private agencies such as Nokia, Natsem and Intel.



Pabitra Mitra

Research Interests: *Machine Learning, Information Retrieval, Data Mining*

Pabitra Mitra did his PhD from Indian Statistical Institute Calcutta in 2003. His research interests are in the fields of machine learning, data mining, information retrieval, and pattern recognition. He has authored a book on Data Mining and about twenty papers in international journals. He is a recipient of the Indian National Academy of Engineering Young Engineer Award in 2007. His hobbies are painting and reading story books.



Jayanta Mukhopadhyay

Research Interests: *Image and Video Processing, Pattern Recognition and Multimedia System.*

Jayanta Mukhopadhyay received his B.Tech., M.Tech., and Ph.D. degrees in Electronics and Electrical Communication Engineering from the Indian Institute of Technology (IIT), Kharagpur in 1985, 1987, and 1990, respectively. He joined the faculty of the Department of Electronics and Electrical Communication Engineering at IIT, Kharagpur in 1990 and later transferred to the Department of Computer Science and Engineering where he is presently a Professor. He served as the head of the Computer and Informatics Center at IIT, Kharagpur from September 2004 to July 2007. He was a Humboldt Research Fellow at the Technical University of Munich in Germany for one year in 2002. He also has held short term visiting positions at the University of California, Santa Barbara, University of Southern California, and the National University of Singapore. His research interests are in image processing, pattern recognition, computer graphics, multimedia systems and medical informatics. He has published over 100 papers in journals and conference proceedings in these areas. He received the Young Scientist Award from the Indian National Science Academy in 1992. Dr. Mukherjee is a Senior Member of the IEEE. He is a fellow of the Indian National Academy of Engineering (INAE).



Debdeep Mukhopadhyay

Research Interests: *Cryptography, Side Channel Analysis, VLSI of Cryptographic Algorithms, Cellular Automata*

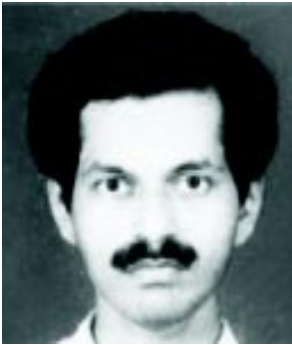
Debdeep Mukhopadhyay is presently working as an Assistant Professor in the Computer Sc and Engg Dept from June 2009. Prior to this he worked as an Assistant Professor in the Dept of Computer Sc and Engg, IIT Madras. Debdeep obtained his BTech from the Dept of Electrical Engg, IIT Kharagpur in 2001. Subsequently he obtained his MS Degree in 2004 and PhD from the Dept of Computer Sc and Engg, IIT Kharagpur in 2007. He has authored about 10 Journal and 49 Conference papers and has served in the Program Committee and as Reviewers of several International Conferences and Journals. Debdeep has been awarded the Indian Semiconductor Association (ISA) TechnoInventor award for best PhD Thesis in 2008.



Ajit Pal

Research Interests: *Embedded Systems, Low-power VLSI Circuits, Sensor Networks and Optical Communication.*

Ajit Pal is currently a Professor in the Department of Computer Science and Engineering at Indian Institute of Technology Kharagpur. He received his M. Tech. and Ph.D. degrees for the Institute of Radio Physics and Electronics, Calcutta University in 1971 and 1976, respectively. Before joining IITKGP in the year 1982, he was with Indian Statistical Institute (ISI), Calcutta, Indian Telephone Industries (ITI), Naini and Defence Electronics Research Laboratory (DLRL), Hyderabad in various capacities. He became full Professor in 1988 and served as Head of Computer Center from 1993 to 1995 and Head of the Computer Science and Engineering Department from 1995 to 1998. His research interests include Embedded Systems, Low-power VLSI Circuits, Sensor Networks and Optical Communication. He is the principal investigator of several Sponsored Research Projects including 'Low Power Circuits' sponsored by Intel, USA. He has over 125 publications in reputed journals and conference proceedings and a book entitled 'Microprocessors: Principles and Applications' published by TMH. He is the Fellow of the IETE, India and Senior Member of the IEEE, USA.



Sudebkumar Prasant Pal

Research Interests: *Design and Analysis of Computer Algorithms, Particularly in the Domain of Geometry and Graph Theory*

Sudebkumar Prasant Pal has research interests in the design and analysis of computer algorithms, particularly in the domain of geometry and graph theory. His current research contributions and interests are in the areas of (i) visibility problems in polygons, (ii) hypergraph coding and coloring, (iii) combinatorial aspects of multipartite quantum entangled states, and (iv) entanglement-assisted quantum protocols defined across a network of remote sites. In the area of computational geometry, he has contributed results on weak and convex visibility, and on the computational and combinatorial complexity of regions visible with multiple specular and diffuse reflections. He has also worked on algorithms for channel routing, and robust high-precision algebraic and geometric computation. In recent years, he has worked on (i) combinatorial characterizations of LOCC incomparable ensembles of multipartite quantum entangled states, and (ii) purely caching based video feeds as opposed to streaming, for scalable video service by introducing the notion of virtual caching in internet proxies. He has held positions such as (i) Convenor, Advisory Committee for the Centre for Theoretical Studies, I.I.T., Kharagpur, and (ii) Member Executive Council: Indian Association for Research in Computing Science. He received the Rajiv Gandhi Research Grant for Innovative Ideas in Science and Technology, 1993, from The Rajiv Gandhi Foundation and Jawaharlal Nehru Centre for Advanced Scientific Research (JNCASR), Jakkur, Bangalore. He worked as Visiting Associate Professor in the Mathematics and Computer Science department in the University of Miami, Florida, USA during the period, August 1999 to May 2000.



Dipanwita Roychowdhury

Research Interests: *Design and Analysis of Cryptographic Algorithms, Theory and Application of Cellular Automata and VLSI Design and Testing*

Dipanwita Roychowdhury is a Professor in the Department of Computer Science and Engineering, Indian Institute of Technology, Kharagpur, India. She received her B.Tech and M.Tech. degrees in Computer Science from University of Kolkata in 1987 and 1989 respectively, and the PhD degree from the department of Computer Science & Engineering, Indian Institute of Technology, Kharagpur, India in 1994. Her current research interests are in the field of Cryptography, Error Correcting Code, Cellular automata and VLSI Design & Testing. She has published more than 125 technical papers in International Journals and Conferences. Dr. Roy Chowdhury has supervised 8 PhD and 6 MS thesis and she is the Principal Investigator of several R&D projects. She is the recipient of INSA Young Scientist Award and Associate of Indian Academy of Science.



Dipankar Sarkar

Research Interests: *Formal Verification and Symbolic Reasoning.*

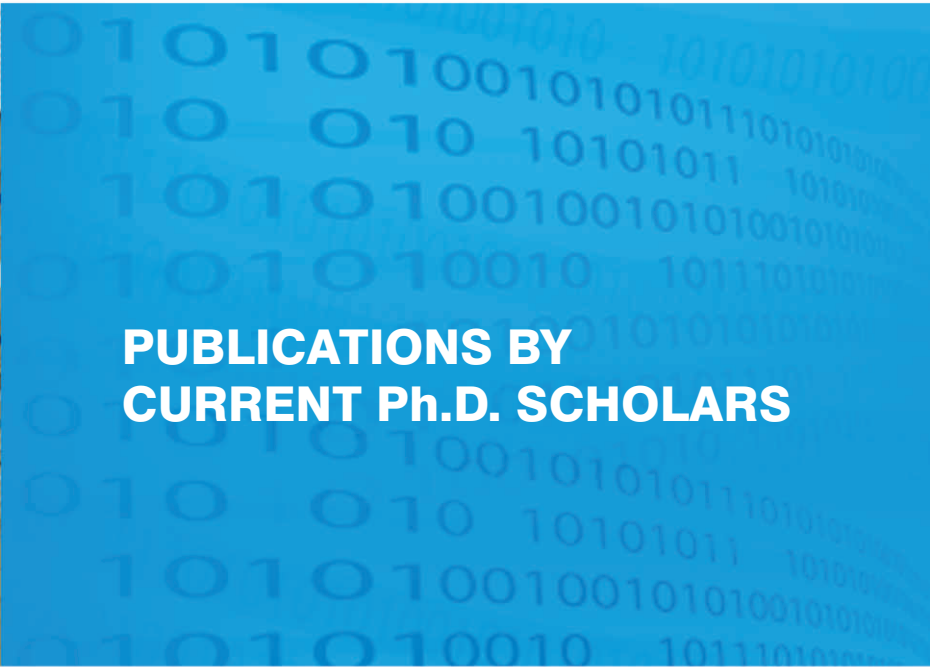
D. Sarkar did his B.Tech., M.Tech. in Electronics and Electrical Communication Engg. and PhD in Engineering from I.I.T., Kharagpur. Has served I.I.T., Kharagpur as a faculty member from 1981.



Sudeshna Sarkar

Research Interests: *Artificial Intelligence, Machine Learning, Information Retrieval, Natural Language Processing*

Sudeshna Sarkar is a Professor in the Department of Computer Science and Engineering at Indian Institute of Technology, Kharagpur. She received the BTech degree in Computer Science & Engineering from IIT Kharagpur, India, in 1989, an MS in Computer Science from University of California, Berkeley in 1991 and Ph.D., in Computer Science & Engineering from IIT Kharagpur in 1996. She has served in the faculty of IIT Guwahati and at IIT Kanpur before joining IIT Kharagpur. Her broad research interests are in Artificial Intelligence and Machine Learning. She is currently working in the fields of natural language processing, text mining and information retrieval and content recommendation systems. She has been a principal investigator in a number of sponsored projects in these areas. Some of these are Cross language information access, Machine Translation between Indian languages, NER and POS tagging, and building of a Bengali treebank. She had been the principal scientist of Minekey, a company incubated at IIT Kharagpur and ran the research centre of Minekey at IIT Kharagpur.



**PUBLICATIONS BY
CURRENT Ph.D. SCHOLARS**

2010

1. B. Mitra, A. K. Dubey, S. Ghose, N. Ganguly, "How do Superpeer Networks Emerge?," *IEEE INFOCOM*.
2. B. Mitra, A. K. Dubey, S. Ghose, N. Ganguly, "Formal Understanding of the Emergence of Superpeer Networks: A Complex Network Approach," *ICDCN*.
3. C. Karfa, D. Sarkar, C Mandal, P. Kumar, "Verification of Datapath and Controller Generation phase in Highlevel Synthesis," *IEEE TCAD*.
4. C. Rebeiro and D. Mukhopadhyay, "Pinpointing Cache Timing Attacks on AES," *VLSID*.
5. P. K. Bhowmick, A. Basu, P. Mitra, A. Prasad, "Sentence Level News Emotion Analysis in Fuzzy Multilabel Classification Framework," *CICLing*.
6. A. Hazra, P. Ghosh, P. Dasgupta, P. P. Chakrabarti, "Coverage Management with Inline Assertions and Formal Test Points," *VLSID*.
7. R. Pal, P Mitra, and J. Mukherjee, "Visual saliency and node centrality measures," *NCVPRIPG*.
8. S. Bag and G. Harit, "A Medial Axis Based Thinning Strategy for Character Images," *NCVPRIPG*.
9. S. K. Dandapat, B. Mitra, N. Ganguly, "Flexible Load Balancing in Wireless Mobile Environment," PhD Forum, *ICDCN* .
10. A. Sarkar, R. Nanda, S. Ghose and P. P. Chakrabarti, SafeERfair "A priori Overload Handling in Fair Scheduled Embedded Systems," *VLSID*.
11. S. Dey, D. Sarkar and A. Basu, "A Tag Machine Based Performance Evaluation Method for Job-Shop Schedules," *TCAD* (Accepted for publication).

2009

1. B. Mitra, "Technological Networks", Edited book volume '*Dynamics on and of Complex Networks: Applications to Biology, Computer Science, Economics, and the Social Sciences*', N. Ganguly, A Deutsch, and A. Mukherjee (eds.) Birkhauser, Springer.
2. C. Rebeiro, D. Mukhopadhyay, Junko Takahashi and Toshinori Fukunaga, "Cache Timing Attacks on Clefia," *INDOCRYPT*.
3. S. S. Ali, C. Rebeiro, and D. Mukhopadhyay, "Cache Aware Tools for Cryptographic Algorithms," *National Workshop on Cryptology*.
4. D. P. Dogra, A. K. Majumdar, S. Sural, "Evaluation of Segmentation Techniques Using Region Size and Boundary Information," *PReMI*.
5. D. P. Dogra, K. Tripathy, A. K. Majumdar, S. Sural, "A Comparative Study on Texture Features Used for Segmentation of Images Rich in Texture," *ICSIPA*.
6. D. P. Dogra, A. K. Majumdar, S. Sural, "Detection of Object Pick Up and Drop Off by Humans in Video Surveillance Applications," *National Seminar on Image Classification and Pattern Recognition*.
7. J. Chandra, N. Ganguly, and S. Shaw, "HPC5: An Efficient Topology Generation Mechanism for Gnutella Networks," *Computer Networks, Elsevier*.
8. J. Chandra, S. Shaw and N. Ganguly, "Analyzing Network Coverage in Unstructured PeertoPeer Networks : A Complex Network Approach," *IFIP Networking*.
9. S. Shaw, J. Chandra, N. Ganguly, "HPC5: An Efficient Topology Generation Mechanism for Gnutella Networks," *ICDCN*.

10. P. K. Bhowmick, A. Basu and P. Mitra, "Reader Perspective Emotion Analysis in Text through Ensemble based MultiLabel Classification Framework," *Computer and Information Science*, Vol. 2(4), pp 6474.
11. P. K. Bhowmick, A. Basu and P. Mitra, "Classifying Emotion in News Sentences: When Machine Classification Meets Human Classification," *International Journal on Computer Science & Engineering*, Vol. 2(1), pp 98108.
12. P. K. Bhowmick, A. Basu, P. Mitra and Abhisek Prasad, "Multilabel Text Classification Approach for Sentence Level News Emotion Analysis," *PReMI*.
13. S. Mitra, P. Ghosh, P. Dasgupta, P. P. Chakrabarti, "Incremental Verification Techniques for an Updated Architectural Specification," *INDICON*.
14. P. Ghosh, S. Mitra, P. Dasgupta, "A Novel Methodology to Assist Client Side Testing of Interactive Web Applications," *ICIT*.
15. P. Ghosh, B. Ramesh, A. Banerjee, P. Dasgupta, "Abstraction Refinement for State Space Partitioning based on Auxiliary State Machines," *IEEE TENCON*.
16. A. Hazra, P. Ghosh, P. Dasgupta, P.P. Chakrabarti, "Inline Assertions Embedding Formal Properties in a Test Bench," *VLSID*.
17. A. Sur, S. Sagar, R. Pal, P. Mitra, and J. Mukhopadhyay, "A new image watermarking scheme using saliency based visual attention model," *IEEE Indicon*.
18. R. Pal, J. Mukherjee, and P. Mitra, "An approach for preparing Groundtruth data and evaluating visual saliency models," *PReMI*.
19. R. R. Suman and R. Mall, "State model extraction of a software component by observing its behavior," *SIGSOFT Software Eng. Notes*, Vol.34(1), pp. 17.
20. S. Chatterji, D. Roy, S. Sarkar and A. Basu, "A Hybrid Approach for Bengali to Hindi Machine Translation," *ICON*.
21. S. Chatterji, P. Sonare, S. Sarkar and D. Roy, "Grammar Driven Rules for Hybrid Bengali Dependency Parsing," *ICON NLP Tools Contest: Indian Language Dependency Parsing*.
22. S. Chatterji, T. M. Sarkar, S. Sarkar and J. Chakraborty, "Karak Relations in Bengali," *AICL*.
23. S. Ghosh, M. Alam, D. R. Chowdhury, and I. Sengupta, "Parallel Cryptodevices for GF(p) Elliptic Curve Multiplication Resistant against Side Channel Attacks," *Computers and Electrical Engineering Elsevier*.
24. M. Alam, S. Ghosh, M jagon Mohan, D. Mukhopadhyay, D. R. Chowdhury and I. Sengupta, "Effect of Glitches Against Masked AES Sbox Implementation and Countermeasure," *Information Security (IFS), IET*, Vol. 3(1), pp 3444.
25. S. Pal, P. Bhowmick, A. Biswas and B. B. Bhattacharya, "Understanding Digital Documents Using Gestalt Properties of Isothetic Components," *International Journal of Digital Library Systems*.
26. S. Pal and P. Bhowmick, "Estimation of Discrete Curvature Based on ChainCode Pairing and Digital Straightness," *ICIP*.
27. S. Pal, P. Bhowmick, A. Biswas, and B. B. Bhattacharya, "GOAL: Towards understanding of Graphic Objects from Architectural to Line drawings," *IAPR Intl. Workshop GREC*.
28. S. Das and D. Roy Chowdhury, "Prevention of Attacks on Grain Using Cellular Automata," *INSCRIPT*.
29. S. Mukherjee and P. Dasgupta, "Incorporating Local Variables in MixedSignal Assertions," *IEEE TENCON*.
30. S. Mukherjee, S. K. Panda and P. Dasgupta, "AssertionBased Verification of MixedSignal Behaviors with Sampling Clock," *SNUG*.

31. S. Mukherjee, A. Ain, S. K. Panda, R. Mukhopadhyay and P. Dasgupta, "A Formal Approach for Specification Drive AMS Behavioral Model Generation," *DATE*.
32. A. Sarkar, S. Swarup, S. Ghose and P. P. Chakrabarti, ERfair "Scheduler with Processor Shutdown." *HiPC*.
33. S. Chatterjee, A. Sarkar and P. P. Chakrabarti. "A priori Overload Detection and Avoidance in RT Fair Scheduled Systems." *HPC ASIA*.

2008

1. B. Mitra, N. Ganguly, S. Ghose and F. Peruani., "Generalized Theory for Node Disruption in Finite Size Complex Networks," *Physical Review E*.
2. B. Mitra, N. Ganguly, S. Ghose and F. Peruani, "Stability Analysis of PeertoPeer Networks Against Churn", *Pramana : Journal of Physics*, Springer, 71.
3. C. Karfa, D. Sarkar, C Mandal, P. Kumar, "An Equivalence Checking Method for Scheduling Verification in Highlevel Synthesis," *In IEEE Transaction on Computer Aided Design on ICs*, Vol 27(3), pp 556-569.
4. Abhigyan, J. Chandra, N. Ganguly, "A Bandwidth Aware Topology Generation Mechanism for PeertoPeer based PublishSubscribe Systems," *ICIIS*.
5. P. K. Bhowmick, A. Mukherjee, A. Banik, P. Mitra and A Basu. "A Comparative Study of the properties of Emotional and Nonemotional Words in Wordnet: A Complex Network Approach," *ICON*.
6. P. K. Bhowmick, P. Mitra and A. Basu, "An Agreement Measure for Determining InterAnnotator Reliability of Human Judgements on Affective Text," *COLING*.
7. R. Pal, P. Mitra, and J. Mukherjee, "Visual saliency based theme and aspect ratio preserving image cropping for small displays," *NCVPRIPG*
8. R. Pal, P. Mitra, and J. Mukhopadhyay, "ICam: maximizes viewers' attention on intended objects," *PCM*.
9. S. K. Saha, S. Chatterji, S. Dandapat, S. Sarkar and P. Mitra, "A Hybrid Approach for Named Entity Recognition in Indian Languages," *IJCNLP Workshop on Named Entity Recognition for South and South East Asian Languages*.
10. S. Ghosh and D. R. Chowdhury, "Elliptic Curve Based Multisignature Scheme for Multiserver Systems," *IEEE TENCON*.
11. S. Ghosh, M. Alam, D. R. Chowdhury and I. Sengupta, "A GF(p) Elliptic Curve Group Operator Resistant Against Side Channel Attacks," *GLSVLSI*.
12. M. Alam, S. Ghosh, D. R. Chowdhury and I. Sengupta, "Single Chip Encryptor/Decryptor Core Implementation of AES Algorithm," *VLSID*.
13. S. Pal and P. Bhowmick, "Cubic Approximation of Curveshaped Objects in Z^2 : A Generalized Approach Based on Discrete Curvature," *PreICM Intl. Convention on Mathematical Sciences*. (Extended version to appear in the Journal of Discrete Mathematical Sciences and Cryptography, Taru Publications, New Delhi, India).
14. S. Das and D. Roychowdhury, "An Efficient $n \times n$ Boolean Mapping Using Additive Cellular Automata," *ACRI*.
15. A. Hazra, A. Banerjee, S. Mitra, P. Dasgupta, P. P. Chakrabarti, C. R. Mohan, "Cohesive Coverage Management for Simulation and Formal Property Verification," *ISVLSI*.

2007

1. B. Mitra, F. Peruani, S. Ghose and N. Ganguly. "Analyzing the Vulnerability of Superpeer Networks Against Attack," *ACM CCS*.
2. B. Mitra, S. Ghose and N. Ganguly, "How Stable are Large Superpeer Networks Against Attack?" *IEEE P2P*.
3. B. Mitra, F. Peruani, S. Ghose and N. Ganguly, "Brief Announcement: Measuring Robustness of Superpeer Topologies," *ACM PODC*.
4. B. Mitra, S. Ghose and N. Ganguly, "Effect of Dynamicity on Peer to Peer Networks", *HiPC* .
5. D. Mandal, S. Nath, B. Mitra, "Survivable Routing in WDM Weighted Network," *IEEE COMSWARE*.
6. H. S. Koppula, K. Puspesh and B. Mitra, "Study and Improvement of Robustness of Existing Networks," *Dynamics On and Of Complex Networks*, A Satellite Workshop of European Conference on Complex Systems.
7. C. Karfa, D. Sarkar, C Mandal, "Verification of Datapath and Controller Generation Phase in Highlevel Synthesis," *ADCOM*.
8. C. Karfa, D. Sarkar and C Mandal, "Handinhand Verification of Highlevel Synthesis," *GLSVLSI*.
9. C. Karfa, C. Mandal, D. Sarkar, C. Reade, "Register Sharing Verification during Datapath Synthesis," *ICCTA*.
10. S. Ghosh, M. Alam, K. Kumar, D. Mukhopadhyay and D. R. Chowdhury, "Preventing the SideChannel Leakage of Masked AES SBox," *ADCOM*.
11. A. Saha and S. Ghosh, "A speedarea optimization of Full Search Block Matching hardware with applications in highdefinition TVs (HDTV)," *HiPC* .
12. S. Ghosh, M. Alam, D. R. Chowdhury, and I. Sengupta, " Effect of Side Channel Attacks on RSA Embedded Devices," *IEEE TENCON*.
13. S. Ghosh and A. Saha, "Speedarea optimized FPGA implementation for Full Search Block Matching," *ICCD*.
14. S. Ghosh, M. Alam, I. Sengupta and D. R. Chowdhury, "A Robust GF(p) Parallel Arithmetic Unit for Public Key Cryptography," *EUROMICRO DSD*.
15. M. Alam, S. Ghosh, D. Mukhopadhyay, D. R. Chowdhury and I.Sengutpa, "Latency Optimized AESRijndael with Flexible Mode of Operation," *VDAT*.
16. A. Saha, S. Ghosh, S. Sural, and J. Mukherjee, "Toward Memoryefficient Design of Video Encoders for Multimedia Applications," *ISVLSI*.
17. M. Alam, S. Ray, D. Mukhopadhyay, S. Ghosh, D. R. Chowdhury, and I. Sengupta, "An Area Optimized Reconfigurable Encryptor for AESRijndael," *DATE*.
18. S. Ghosh, D. R. Chowdhury, and I. Sengupta, "Side Channel Attacks on RSA and ECC Crypto Devices," *National Workshop on Cryptology*.

2006

1. B. Mitra, Md. M. Afaque, S. Ghose, N. Ganguly, "Developing Analytical Framework to Measure Robustness of Peer to Peer Networks," *ICDCN*.
2. B. Mitra, S. Bhattacharjee "Worst Case Delay Minimization in WDM Linear and Ring Networks," *EAIT* .
3. C. Karfa, C. Mandal, D. Sarkar, S. R. Pentakota, C. Reade, "A Formal Verification Method of Scheduling in Highlevel Synthesis," *ISQED*.
4. S. Biswas, C. Karfa, H. Kanwar, D.Sarkar, S. Mukhopadhyay A. Patra, "Fairness of Transitions in Diagnosability Analysis of Hybrid Systems," *ACC*.

5. C. Karfa, C. Mandal, D. Sarkar, S. R. Pentakota, C. Reade, "Verification of Scheduling in Highlevel Synthesis," *ISVLSI*.
6. D. P. Dogra, "A Hidden Markov Model Based Approach for Telephonic Digit Recognition," *EAIT*.
7. A. Sarkar, P. P. Chakrabarti and R. Kumar. "FrameBased Proportional RoundRobin," Published in: *IEEE Transactions on Computers*, Vol. 55(9), pp 1121-1129.
8. A. Sarkar, P. P. Chakrabarti and R. Kumar. "FrameBased Fair Multiprocessor Scheduler: A Fast Fair Algorithm for Realtime Embedded Systems." *VLSID*.

2005

1. D. Mandal, B. Mitra "Shared Path Protection in DWDM Mesh Networks," *CIT* .
2. B. Mitra, S. Nath "Design of a Survivable Virtual Topology for DWDM Networks," *CSI Conference on Networks*.
3. C. Karfa, J. S. Reddy, S. Biswas, C. R. Mandal, D. Sarkar, "SAST: An Interconnection aware high level synthesis tool," *VDAT*.
4. D. P. Dogra, H. Karmakar, "Realtime Uncompressed Voice Transmission Through High Speed LAN," *National Conference on Applications of Advanced Technology in Networking*.
5. A. Sarkar, P. P. Chakrabarti and R. Kumar. "BoundaryFair RoundRobin: A Fast Fair Scheduler." *VDAT*.

2004

1. R. Datta, B. Mitra, S. Ghose, I. Sengupta "An Algorithm for Optimal Assignment of a Wavelength in a Tree Topology and Its Application in WDM Networks" *IEEE Journal on Selected Areas in Communications*, Vol. 22(9), pp 1589 - 1600.
2. B. Mitra, A. K. Mandal "Wavelength Assignment in Optical Line Networks and Its Application in WDM Ring Topology", *MobiComNet*.
3. R. Ghose, S. Kar, B. Mitra and S. Ghose, "Physical Topology Design For All Optical WDM Networks Under LAMBDA Constraint", *CODEC*.
4. S. P. Pal, R. R. Suman, G. S. Anil Kumar and R. Malhotra, "Virtual Video Caching: A scalable & generic technique for improved quality of video service," *Journal of HiPC*, Vol. 13(4), pp 249-263.

2003

1. S. P. Pal, R. R. Suman, G. S. Anil Kumar and R. Malhotra, "Virtual Video Caching: A scalable & generic technique for improved quality of video service," *Trusted Internet Workshop*.

