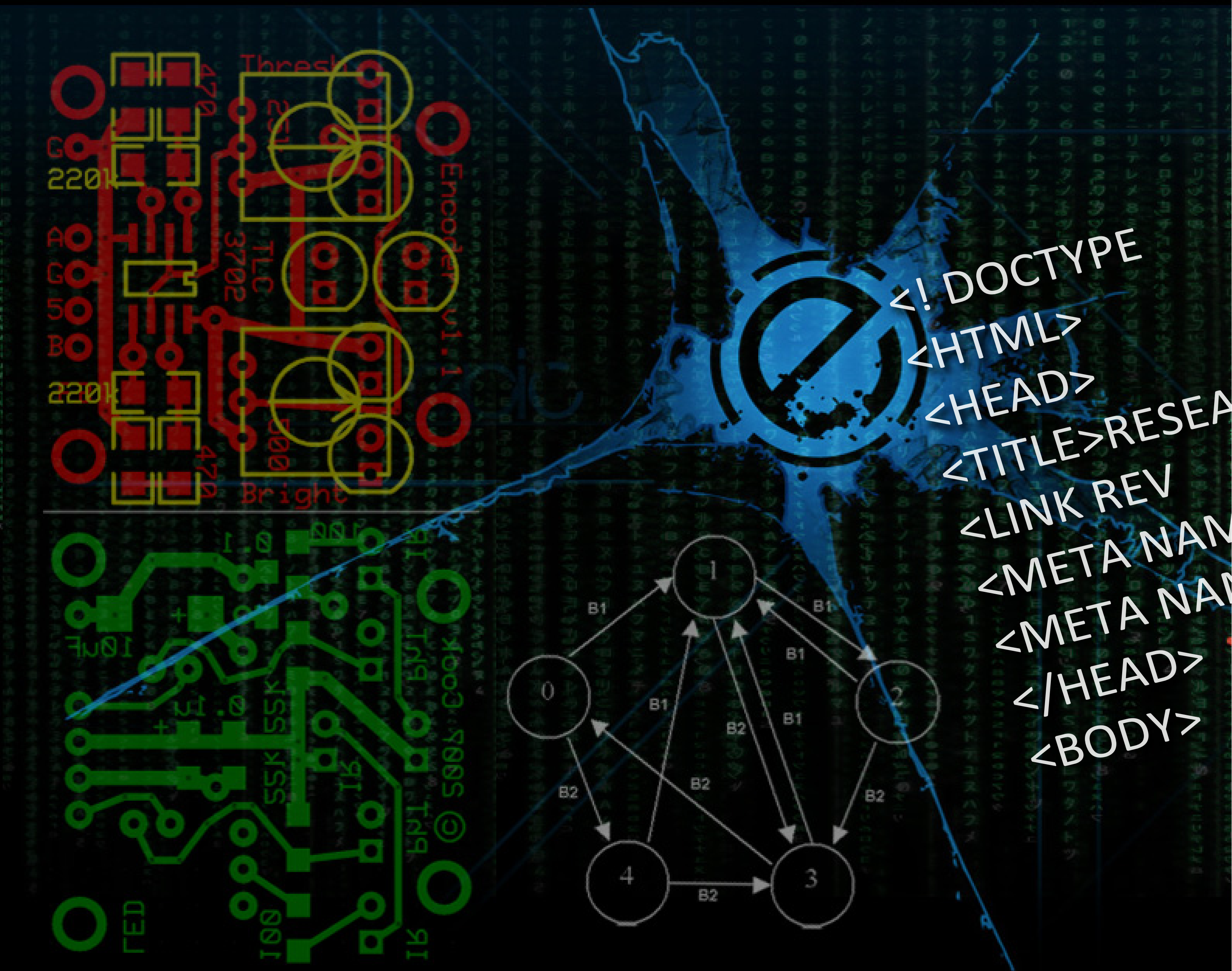
Research Scholars' Day

Department of Computer Science & Engineering

February 12, 2011



Indian Institute of Technology, Kharagpur

Department of Computer Science and 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 centers 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 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

The grand success of the first Research Scholar's Day held last year has been a matter of great pleasure and satisfaction for us, and now we stand on the doorstep of the second research scholar's day on 12th February, 2011. We look forward to observing this day with zeal and enthusiasm.

The event is aimed to provide a forum for demonstrating the latest research findings and exchanging ideas of research, development, and knowledge among the different research groups and faculties of the department. The one-day gathering will consist of short presentations by the research scholars, poster presentations along with a panel discussion, and a faculty-student interaction session.

I note with satisfaction the collective effort put in by the students and faculty of the department for this event. I wish them all the success.

Jayanta Mukhopadhyay

PhD Scholars

List of Current Ph.D. Scholars

Aritra Hazra
Arnab Sarkar
Bibhas Ghosal
Bodhisatwa Mazumdar
Chandan Karfa
Chhabi Rani Panigrahi
Chester Rebeiro
Debi Prosad Dogra
Dinesh Dash
Ishani Chakraborty
Joydeep Chandra
Kallol Mallick
Kamalesh Ghosh
Kunal Banerjee
Mahesh Raghunath Shirole
Manjira Sinha
Maunendra Sankar Desarkar
Paranatapa Bhattacharya
Prasenjit Mondal
Priyankar Ghosh
Rajeev Ranjan Suman
Rajendra Prasath R
Rajib Ranjan Maiti
Rishiraj Saha Roy
Ruchira Naskar

Sabyasachi Karati
Sandip Karmakar
Sanjay Chatterjee
Sankara Reddy D
Santosh Ghosh
Saptarshi Ghosh
Satya Gautam Vadlamudi
Shyamosree Pal
Sk Subidh Ali
Soma Saha
Soumen Bag
Soumyadip Bandyopadhyay
Soumyajit Dey
Sourav Kumar Dandapat
Srinivasa Rao Myla
Srobona Mitra
Subhadip Kundu
Subhankar Mukherjee
Subhasish Dhal
Subhendu Bhadra
Subrata Nandi
Sudip Roy
Sudipta Saha
Sumanta Pyne
Tirthankar Dasgupta

Research Abstracts
(PhD)

Aritra Hazra
Email: aritrah@cse.iitkgp.ernet.in
Joined the department in: July 2010

***Aritra Hazra** received a B.E. degree in Computer Science and Engineering from Jadavpur University, Kolkata in 2006, and an M.S. degree in Computer Science and Engineering from Indian Institute of Technology, Kharagpur in 2010. From July 2006, he worked in several projects of SRIC, IIT Kharagpur, as a Research Consultant. The projects are primarily in the following fields: Design Intent Verification and Coverage Analysis, Power Intent Verification of Power-managed Designs, Platform Architecture Modeling for Exploring Power Management Policies. Since July 2010, he has been a research scholar in the department of Computer Science & Engineering in IIT Kharagpur. His research interests are in the areas of Design Verification, Power Intent Verification, Assessment and Improvement of Functional Reliability.*

Supervisors: Prof. Pallab Dasgupta and Prof. Partha Pratim Chakrabarti

Verification and Coverage of Architectural Power Intent

The battle against increasing power-densities within chips has also led to an awareness of the need for better architectural power management. Consequently, the task of developing efficient on-chip power management strategies has been pushed up to the architectural level and has become a significant challenge for the micro-architects. Several tools have been developed for early estimation of power performance, exploring alternative strategies and converging on the most efficient one. Power performance analysis typically leads to the development of a global power management strategy, which demarcates the boundaries of various architectural power domains and specifies the properties relating these power domains at a high level of abstraction. The architectural power management strategy may entail simple properties (such as domain-A and domain-B will not be active simultaneously) or complicated properties, such as ones which specify start-up sequences between the power domains in a complex system-on-chip (SoC).

Though global power management strategies are designed up front, they can be implemented only much later in the design flow. This is because, power intent specification is not supported in high-level design languages such as Verilog or SystemVerilog which are used to enter the design implementation at behavioral or Register-Transfer Level (RTL). Typically, both the design and the power control logic (PCL) can be expressed in digital logic and can be coded using hardware description languages (such as Verilog/VHDL). Since the power lines, voltage regulators or level-shifter circuits cannot be described in RTL, it is not possible to directly express the power control circuitry (PCC) in RTL. The adoption of new standards like UPF (Unified Power Format) and CPF

(Common Power Format) for specifying PCC has led to new opportunities for accelerating the verification of power-managed designs. With the help of suitable UPF constructs, the power architects can specify the power domains supporting multiple voltages, frequencies and body-biasing features – thereby introducing more rigor in the verification of PCL.

Assertions extracted from UPF specifications typically target the verification of individual local power managers (LPMs) and are per-domain in nature. On the other hand, properties meant for the global power manager (GPM) are typically inter-domain in nature and are decided top-down, often at the architectural level. The architectural power intent of a design typically expresses high-level properties relating multiple power domains without specific reference to the low-level signals which control the power states of the individual domains. Recent research has indicated ways of using UPF specifications for extracting valid low-level control sequences to express the transitions between the power states of individual domains. Today, there is a disconnect between the high-level architectural power management strategy which relates multiple power domains and these low-level assertions for controlling individual power domains. In our work, we attempt to bridge this disconnect by leveraging the low-level per-domain assertions for translating architectural power intent properties into global assertions over low-level signals.

We propose an approach where architectural power intent properties are expressed formally using several pre-defined predicates related to abstract interpretations of the states of the architectural power domains. We, then, provide a methodology for automatic translation of these architectural properties into assertions using the low-level signals (which may come into existence much later in the design flow). Our translation leverages the per-domain properties extracted from UPF specifications. We show that the inter-domain properties, created in this manner to mimic the architectural power intent, can be formally verified over the global power management logic using an industrial formal verifier. Similarly, the per-domain assertions, extracted directly from UPF for individual power domains, can also be verified over the power management logic for the individual domain. Moreover, verification of a power management unit involves verifying the global power manager in multiple legal power states and verifying only the intended transitions and sequences of transitions have occurred. Therefore, measuring the functional coverage of the power management unit with respect to the architectural power intent is a significant challenge. In future, we shall try to extend this work to propose a framework for indicating the simulation-based verification coverage of the power manager by monitoring the UPF-extracted per-domain assertions as well as the generated global assertions while simulating the power control logic with suitable test benches. Similarly, we shall try to explore a methodology to indicate the formal verification coverage of global power management unit introducing suitable coverage metrics. We believe that this work provides a new direction for the verification and coverage of inter-domain power management logic against the architectural power management strategy.

Arnab Sarkar
Email: arnab@cse.iitkgp.ernet.in
Joined the department in: July 2006

__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 to 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 and *Partition Oriented ERfair Scheduler (POES)*, a partition-merge based algorithm that reduces both the number of migrations and cache-misses by upto 40 times. Our algorithm *ERfair Scheduler with Shutdown on Multiprocessors (ESSM)* attempts to reduce processor energy consumption in ERfair systems by locally maximizing processor shut-down intervals through a novel procrastination scheme. We have also developed overload handling strategies for Pfair and ERfair multi-processor systems using efficient low-overhead admission control. 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.

Bibhas Ghoshal
Email: bibhas@cse.iitkgp.ernet.in
Joined the department in: December 2009

Bibhas Ghoshal *is pursuing his PhD in Computer Science and Engineering from II, Kharagpur. His recent research activities have been in the areas of VLSI testing with focus on testing of Network-on-chip architectures. He holds an ME (2005) degree in Computer Science and Engineering from West Bengal University of Technology and M.Sc (2002) degree in Electronic Science from Jadavpur University.*

Supervisor: Prof. Indranil Sengupta

New Approaches in Testing of Network-on-chip Architectures

As System-on-chips (SOCs) are getting packed with multiple-cores, they are becoming more and more memory dominant. Thus, the testing of memory cores contributes significantly in the testing time of SOCs. Research suggests that, two most effective ways of test time reduction of SOCs is reusing an already existing on-chip intra-core communication network, Network-on-chip (NOC) as Test Access Mechanism (TAM) and using a Built-in-self-Test(BIST) technique to test the individual memory cores. However, memory cores with BIST structure increase the area overhead, if the number of memory cores is large. BIST cores also imply large test patterns and hence increase the test time. We have proposed a hierarchial BIST structure as a solution to the memory core test problem. The structure consists of number of BISTed memory cores and BIST core controllers interconnected via Network-on-chip communication platform. All memory cores in SOC based systems are not of the same type. They vary according to different parameters like size and process technology. Thus in a system with number of memory cores with different parameters, the cores must be divided into groups and each group must be treated accordingly. The structure of SOC thus proposed is a cluster of memory groups with a BIST controller for each group interconnected by Network-on-chip platform. The controllers are controlled by a top level controller. The top level controller sends high level test instructions for the other lower level BIST controllers which pass these instructions to the BISTed memory cores. The memory cores have their own BIST circuit for generating the test data. The cores report the result of test back to the controllers via the same NOC. Sharing of hardware reduces the area overhead problem and the reuse of NOC removes the routing constraint due to hardware sharing. On chip test pattern generation for the cores reduces the test time. Since only high level instructions and final fault report are communicated through the NOC, the communication latency is also substantially low which in turn contributes to the effect of low test time.

Bodhisatwa Mazumdar
Email: bm.iitkgp@gmail.com
Joined the department in: July 2009

***Bodhisatwa Mazumdar** received his B.Tech. degree in Electronics and Instrumentation Engg. from University of Kalyani, Kalyani in 2004, and an M.S. degree in Electronics and Electrical Communication Engg from Indian Institute Institute of Technology, Kharagpur in 2007. From September 2007 till May 2008, he worked in GE Healthcare, Bangalore, as a Hardware Design Engineer. Since May 2008 to July 2009 he worked as Member Technical Staff in Manthan Semiconductors Pvt. Ltd., Bangalore. Since July 2009, he has been a research scholar in the Department of Computer Science & Engineering in IIT Kharagpur. His research interests are in the areas of Cryptography and Network Security.*

Supervisors: Prof. Debdeep Mukhopadhyay and Prof. Indranil Sengupta

Power Analysis Attacks on Stream Cipher Cryptosystems

In modern cryptography, stream ciphers are most useful in applications where information needs to be processed at high speed and resource availability is a constraint. In literature, there exist plenty of stream ciphers whose internal states are based on Linear Feedback Shift Registers (LFSR) and that they use modular additions to produce output streams. The abundance of these LFSR-based stream ciphers with modular additions can be attributed to the fact that, when they are efficiently implemented in hardware, they produce outputs with high throughput. The security of stream ciphers has only recently received increasing attention.

Till recent times, research towards security of cryptographic algorithms has been concerned with classical cryptanalysis. In reality, an adversary can mount an attack on the implementation of the algorithm even if the cipher stands on an abstract mathematical model which has been intensely investigated. Such attacks called Side-channel Analysis (SCA) can use all kinds of physical emanation from the device, like power consumption, electromagnetic radiation, or execution time variations. The objective of the present research work is to assess the strength of stream cipher cryptosystems based on non-linear filter (NLF) generators against differential power analysis (DPA) attacks and to develop appropriate countermeasures against them.

The scope of current research work includes developing DPA attacks on stream ciphers based on nonlinear filter generator, evaluate the current eSTREAM portfolio stream ciphers for e.g. Grain v1, and develop countermeasures for the DPA attacks on the above-mentioned stream ciphers.

The work that have been done so far includes literature review of different modes of power analysis attacks on stream ciphers on eSTREAM project (the ECRYPT stream cipher project) portfolio for e.g. Grain v1, Trivium. Also experiments were carried out to determine the variation of power dissipation for stream cipher execution with respect to hamming distance of the LFSR states of the stream cipher over a number of clock cycles on Xilinx XPower tool. Moreover, on this power analysis based SCA technique, the work focusses to validate the claim that the state of an n-bit LFSR can be determined by measuring the power consumed by the LFSR in each cycle over consecutive cycles. The experiments revealed that the power profile of the stream cipher has a strong correlation with the hamming distance of the LFSR states. Numerous parity equations have been generated from the LFSR feedback limited by the LFSR sequence length. At present we attempt to extract LFSR hamming distance profile from the power profile through probabilistic analysis of satisfying these parity equations by the bit sequence obtained from the average power measurements on consecutive clock cycles after thresholding operation on the average power values.

Chandan Karfa
Email: ckarfa@cse.iitkgp.ernet.in, ckarfa@yahoo.co.in
Joined the department in: Januray 2008

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. Chittaranjan Mandal and Prof. Dipankar Sarkar

Equivalence Checking in Embedded System Design Verification

Present day embedded systems synthesis consists in application of several sophisticated transformation techniques on the input behaviours 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 behavioural verification; it shows that executions depicted by the input behaviour 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 apporaches to verify them are briefly discussed.

Several speculative code motion transformations, arithmetic transformations such as, renaming, common sub-expressions elimination and algebraic transformations, etc, may be applied on the input sequential behaviours at the preprocessing stage of embedded system synthesis. The input behaviours are transformed significantly due to these transformations. We have developed an FSMD (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, our method is strong enough to handle both uniform and non-uniform types of code motions, several arithmetic transformations and the cases of control structure modifications of the original behaviors.

A high-level behaviour is mapped to register transfer level (RTL) description during the behavioural synthesis process. The RTL consists of a description of the datapath netlist and a controller FSM. Towards establishing equivalence between a given high-level behaviour and its corresponding RTL behaviour, we have developed a rewriting based method to extract the high-level behaviour from the RTL description and then apply our FSMD 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.

Loop transformations techniques are another set of transformations that are applied extensively on array and loop intensive behaviours in the design of area/energy efficient systems in the domain of multimedia and signal processing applications. Ensuring correctness of these transformations is crucial for the reliability of the systems. We have developed an array data dependence graph (ADDG) based equivalence checking method for this problem. Our method is can handle several arithmetic transformations along with several kinds of loop transformations. At present, we are developing an equivalence checking method by combining the features of both the FSMD based and ADDG based methods.

To deploy a given sequential behavioural code on a highly concurrent heterogeneous multi-processor system, the given code is often converted to a parallellised version. In the case of streaming applications, the Kahn process network (KPN) is commonly used to model the parallel behaviour. Equivalence between the original sequential behaviour and its corresponding KPN behaviour must be insured. In this work, we model both the sequential and the KPN behaviours as ADDGs and then apply our ADDG based method to establish the equivalence.

The above parrallising mechanism usually extracts the maximal concurrency which may be supported by available implementation platform. In that situattion, the KPN obtained will have to be modified to adjust the level of concurrency to match the target platform. Transformation techniques such as, process splitting, channel merging, process clustering, and unfolding and skewing, may be used for this purpose. While direct checking with the original sequential behaviour is an option, it may be desirable to establish direct equivalence between two KPN behaviours (on account of inherent incompleteness of the equivalence checking method). We plan to extend our ADDG based method to this effect.

Chhabi Rani Panigrahi
Email: Chhabi@cse.iitkgp.ernet.in
Joined the department in: July 2009

Chhabi Rani Panigrahi *received her Master in Computer Application from Berhampur University, Orissa in 2000 and an M.Tech. degree in Computer Science and Engineering from AAI Deemed University, Allahabad in 2007. She worked in RS software Pvt. Ltd., Bangalore as a software development engineer from August 2000 till June 2001. From July 2001 to till date, she has been working as a Senior Lecturer in the Department of Computer Science and Engineering at Seemanta Engineering College (under BPUT), Orissa. Since July 2009, she has been a research scholar in the department of Computer Science & Engineering in IIT Kharagpur, as a sponsored candidate under QIP scheme. Her research interests are in the areas of regression testing of object-oriented programs.*

Supervisor: Prof. Rajib Mall

Model-Based Regression Test Case Prioritization

Whenever a program is modified, regression testing is carried out to ensure that the unmodified features of the program continue to work satisfactorily. Regression testing is an expensive activity and typically accounts for as much as half the total software maintenance costs. To reduce the regression test effort, various approaches such as regression test selection (RTS), test suite minimization (TSM) and test case prioritization (TCP) have been proposed in the literature. These techniques target to reduce the cost and improve the effectiveness of regression testing.

An important limitation of both RTS and TSM approaches is that they do not impose any ordering among the selected test cases. As a result, these approaches do not provide the tester with the option of choosing the best subsets of test cases for regression testing under specified cost or time constraints. In this context, TCP techniques target to order regression test cases to increase the rate of fault detection or maximize the rate of code coverage. The existing TCP techniques do not work satisfactorily for TCP of object-oriented programs, because these approaches do not consider the implicit dependencies that arise due to object-relations.

Model-based analysis has several advantages. The algorithms available for graph analysis are usually much more efficient than those for textual analysis of program code. In this context, we propose a model-based approach to prioritize regression test cases for object-oriented programs. Our proposed graphical model represents several relevant object-oriented features such as inheritance, polymorphism, association, aggregation and exceptions. We construct a forward slice of the model to identify all the model elements that may be affected by a change. To determine all model elements indirectly tested by the test case, we construct backward slices with respect to each model element executed by a test case. All the affected model elements and the elements being tested are used to prioritize test cases.

Further, in our model we have included dynamic aspect of object-oriented programs, that is, message path sequences from UML sequence diagrams that cannot easily be analyzed from source code. This information also helps in reducing the size of the constructed slices and helps to increase the accuracy of our TCP technique. In our prioritization approach, we also consider the dependencies among test cases because in many practical situations there exist dependencies among test cases. For example, in a *Library System*, a test case for *Issue Book* use case cannot be executed unless the test cases for *Create Book* use case have been executed successfully.

Chester Rebeiro
Email: chester@cse.iitkgp.ernet.in
Joined the department in: July 2009

***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 and Cryptanalysis, Computer Architecture, and VLSI*

Supervisor: Prof. Debdeep Mukhopadhyay

Cache Based Side Channel Analysis of Cryptographic Algorithms

With the advent of e-commerce and electronic transactions, the need for development of secured systems has grown tremendously. However history has taught us that designing strong cryptographic algorithms is just the beginning. Although the internal algorithms are robust against conventional cryptanalysis, implementations of the ciphers may leak valuable secret information through covert channels commonly referred to as *side channels*. Cache-attacks are a class of side-channel attacks that affect all systems that use cache memories. These attacks glean secret information from the cache access patterns. The cache patterns are manifested through the power consumed and the time required for an encryption. This research work aims at understanding cache-attacks with the aim of defending against the attack.

Cache-attacks affect every system that uses cache memories. Unlike other forms of side channel attacks, cache-attacks do not always require a close proximity to the encrypting device. In 2009, Crosby, Wallach, and Riedi showed that it is possible to gather timing information about encryptions from any computer connected on the Internet. Thus any computer on the Internet can be compromised by this form of attack. Due to this threat, microprocessor manufactures like Intel have recently added dedicated instructions to prevent cache attacks. However these instructions are only useful for the Advanced Encryption Standard (AES). All remaining ciphers in use today are still vulnerable to being attacked by leakages through the cache.
Current defense strategies are mostly in the implementation of the cipher. These strategies are either infeasible in practice or have an enormous computation time overhead making the cipher unsuitable for most applications. Further there is also no measure of the level of security provided against such

counter-measures. Thus it is possible that a countermeasure for one attack would open a vent for another. Alternate strategies require an over hauling of the cache architecture. In 2007, Lee and Wang found that cache interference was the root cause and proposed two designs: a partition locked cache (PLcache) and random permutation cache (RPcache). However incorporating these new cache designs on existing systems is not feasible. Formal models have been developed to counter side-channel attacks. Also, a few provably side-channel resistant ciphers have been proposed. However these proposals are often complex and does not lead to efficient design implementations. In some cases it has also been found to be still vulnerable. This shows that there is a significant gap between the theory developed and the practical world.

The research explores cache attacks with the aim of developing systems which are secure against this deadly attack. First the power of cache-attacks is established by minimizing the attack time and expanding the configurations in which cache-attacks are possible. Minimizing the attack time would require more optimized attacks. One technique is to combine classical cryptanalytic techniques with cache attacks. A cache attack enhanced with differential cryptanalytic techniques can reduce the complexity of an attack on the block cipher CLEFIA from 2^{40} to 2^{14}. In order to determine the kinds of ciphers susceptible to such attacks, ciphers with small table implementations were targeted. A profiled cache timing attack on CLEFIA implemented with small tables requires around 2^{26} encryptions and accurately obtains the secret key 85% of the time. The attack time is less than 5 minutes.

The next part of the research is to pin-point the main causes of cache attack. That is the features in the cache memory and cipher structure most susceptible to cache attacks are identified. It was found that in AES; only the first two rounds are vulnerable. Thus only the first two rounds need to be protected. Similarly, it was found that cache features such as prefetching, parallelization, pipelining, and out-of-order execution lead to attacks.

Cache-attacks would subsequently be formally modeled in order to quantify the amount of information leaked through the cache memory. These models would be used to develop a metric, which could be used to quantify the security of a cipher and its implementation. This would aid in quantifying the security offered by counter-measures and help a naive programmer who is ignorant of the leakages in cache memory.

Debi Prosad Dogra
Email: dpdogra@cse.iitkgp.ernet.in
Joined the department in: July 2007

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 & Engineering in IIT Kharagpur. His research interests are in the areas of Image and Video Processing.*

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

Segmentation and Tracking Algorithms for Hammersmith Infant Neurological Examination

Improving health care services by means of information technology is being considered as a pragmatic approach for delivering improved services toward human community. Various information technology based solutions exist for smoothly running social and medical services. Automation of health care services is mainly driven by efficient computational and interdisciplinary technologies. For example, usage of tele-medicine based systems in hospitals, health centers and remote places are common. Both, information technology and advanced communication technology have made it possible. In today’s scenario, high speed data transfer, efficient computational power and sophisticated medical instruments are the backbone of such health care revolutions. In western countries, it has been in use successfully for a considerable amount of time and now in India, we aim to reach such a state so that more people can be benefited.

Medical imaging and signal analysis techniques like Magnetic Resonance Imaging (MRI), Ultrasound Imaging (USG), X-Ray, Computed Tomography (CT) Imaging, Electrocardiography (ECG) and Electroencephalography (EEG) are widely used for diagnosing patients. Automatic analysis of the signals received through the above mentioned modalities can decrease the turnaround time of examinations without degrading the quality of outputs. But, these imaging modalities are mainly used to understand the state of the internal structure or functionality of various human body parts and organs. Similarly, external imaging of human body using camera is used in related examinations. For example, understanding human gait, learning aerobics and treating mentally ill persons require analysis of images and videos of human activities in front of camera.

Advancement in perinatal care through intensive medical care performed in Neonatal Intensive Care Units (NICU) has increased the survival rate of very low birth weight (VLBW) and preterm newborns. An early prediction based on the outcome of various examinations conducted while the infant is admitted in NICU is clinically useful to counsel families who may benefit from early interventions. However, according to the experts, a significant proportion of highly preterm infants (gestational age at birth less than 32 weeks) usually show neurological and developmental disabilities. For example, due to Intra-Ventricular Haemorrhage (IVH) or Peri-Ventricular Leukomalacia (PVL), an increasing prevalence of Cerebral Palsy (CP) can occur in premature, low birth weight (LBW) newborns and children born with asphyxia. But, an early diagnosis of such disorder can increase the survival rate. A method to diagnose neurological disorder is known as Hammersmith Infant Neurological Examination (HINE). It is a quantitative method for assessing neurological development of infants between 2 and 24 months of gestational age. The examinations include assessment of cranial nerve functions, posture, movements, tone, reflexes, etc. While examinations are being carried out, 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.

However, we have not found any application that tries to automate the HINE examination process. Therefore, till date, doctors conduct and record the outcomes of HINE manually. This has motivated us to propose a scheme using the visual evidence of the examination for achieving speedup and accuracy in the process. In order to have better estimation of neurological parameters, video and image processing based algorithms can be developed to facilitate estimation of parameters like adductors, popliteal angle, ankle dorsiflexion, arm protection, pulled-to-sit, etc. We have proposed algorithms for semi-automatic scoring for a number of examinations which can be modeled using images of the scene. On the other hand, a novel tracking algorithm has been proposed which is found to be successful in automating various video based examinations. An interface for patient care management has also been developed. The feedbacks we are receiving from the doctors of hospitals are encouraging. Medical community is becoming interested in using such health care systems to expedite various examination processes.

Dinesh Dash
Email: dd.dineshdash@gmail.com
Joined the department in: January 2008

Dinesh Dash *received M.Sc. degree in Computer & Information Science from University of Calcutta, Kolkata in 2002, and M.Tech. degree in Computer Science from the same university in 2004. From July 2004 to June 2007, he worked in Asansol Engineering college, Asansol under the West Bengal University of Technologies, India as a Lecturer. Since January 2008, he has been a research scholar in the department of Computer Science & Engineering in IIT Kharagpur. His research interests are in the areas of geometric applications in Wireless Sensor Network.*

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

Coverage problem in wireless sensor network

There are various applications of sensor network like environment monitoring, forest fire detection, wild life habitat monitoring, health monitoring etc. Sensor has the power to sense temperature, sound, pressure etc. It can make some local computation and can communicate with its neighbor sensors or to a base station. By coverage we mean how efficient the sensor network can detect the sensing event. Providing good coverage is one of the major issues in sensor network. If a sensor network does not provide the coverage as per requirements, the use of sensors is useless. Sensor has some limitation like battery power, computation power, sensing capabilities and a specific communication range. Our object is to design efficient algorithm so that it runs on the sensors and ensure coverage. There are various definitions of coverage metric depending upon its application. *Area coverage* [1], here it is ensured that every point of the selected region is under the sensing range of some specific number of sensors. *Target coverage* [2], in this application there is a set of target points; it ensures that all the target points are in the sensing range of some sensors. *Barrier coverage* [3], here there is a given bounded region whenever you passing through the boundary you have to cross some specific number of sensors' sensing regions.

Object tracking is one of the applications of sensor network. Here the sensors detect a moving object and return its position. Suppose we have a given bounded region and an initial point I and a final point F. We want to track the moving object to its entire path when it goes from I to F. *Maximal support path* is one of the best path for tracking the moving object. If the moving object follows the *maximal support* path between I and F then the farthest distance to the closest sensor is minimum. In one of our work, we ensure the support value of the maximal support path after failure of sensors.

In our second work, we assume the moving objects are walking in straight line path. A line segment is said to be *k-line covered* if it intersects k sensors' sensing regions. We assume that the sensing regions of sensors are circular disk. We provide a deployment scheme to ensure *k-line coverage* for the set of line segments. The deployment scheme is such that it uses minimum sensors to achieve the desired coverage.

We have seen that achieving line coverage using minimum number of sensors is NP-hard. We provide good constant factor approximation as well as PTAS for this problem for a special case where line segments are horizontal or vertical.

In our third work, we have defined metrics called a *k-d free region, k-d covered region*. A region is said to be *k-d free* if all line segments of length less than *d* are never sensed by k sensors. Similarly, a region is said to be *k-d covered* if all line segments of length greater than *d* are always sensed by at least *k* sensors.

.

We have proposed algorithms for determining the value of *d* for a given *k* in a bounded region for a given deployment.

References

1. Chi-Fu Huang, Yu-Chee Tseng; the Coverage Problem in Wireless Sensor Network; Mobile Network and Applications; Vol. 10, No. 4, 2005
2. Maggie X. Cheng, Lu Ruan, Weili Wu; Achieving Minimum Coverage Breach under Bandwidth Constraints in Wireless Sensor Networks; InfoCom , 2005
3. Santosh Kumar, Ten H. Lai, Anish Arora; Barrier Coverage With Wireless Sensors; MobiCom 2005.

Ishani Chakraborty
Email: ishani.chakrab@gmail.com
Joined the department in: January 2009

I am presently in the PHD program in Computer Science Department in IIT, Kharagpur working on faceted text retrieval.I have done MS in Information & Computer Science from University of California, irvine, USA in 2007 and worked in Capgemini U.S. LLC. from 2007 to 2008. I have done MCA in 2002 from BIT Mesra,Ranchi,India. I love Photography and Indian Classical Music.

Supervisors: Prof. Anupam Basu and Prof. Sudeshna Sarkar

Faceted Blog Distillation

Blogging is sometimes viewed as a new, grassroots form of journalism and a way to shape democracy outside the mass media and conventional party politics.Blog sites devoted to politics and punditry, as well as to sharing technical developments, receive thousands of hits a day. But the vast majority of blogs are written by ordinary people for much smaller audiences.These blogs are used as a diary or daily journal, a community forum,a commentary to express opinion, a catharsis to let out one's thoughts and feelings; the motivation for writing blogs is varied.

Since there is so much variation in blog genre and style (originating from the variation in the writers motivation), more than topic-related search will be required to satisfy the user. Faceted search enables us to look at the data from its different facets or faces. For blogs, faceted retrieval is not limited to topics but retrieves blogs based on their opinionated-ness, genre, writing style and many others.

My work consists of a baseline retrieval subsystem which performs the baseline blog distillation according to topic,an opinion identification subsystem,an in-depth analysis subsystem and personal blog identification subsystem which performs the faceted blog distillation task. In the baseline system, documents which are deemed relevant are retrieved by the retrieval system with respect to the query, without taking into consideration of any facet requirements.A new weighting model is introduced in the baseline retrieval system which weighs those words highly which are present in lesser number of documents with greater density.In the opinionated vs. factual and personal vs. official faceted retrieval,the results obtained in baseline retrieval is post processed based on features selected and then scored and ranked according to the facet.In the in-depth vs. shallow faceted task, the depth of the text is measured by coherence and coverage of the topic and then these are scored and ranked according to score.

Joydeep Chandra
Email: joydeep@cse.iitkgp.ernet.in
Joined the department in: May 2008

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 if 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

Topological Effects on the Performance of Peer-to-Peer Networks

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, a major issuethat needs to be considered is the overlay or topology formation. Topology plays an important role in determining the search performance of the peers in the network and hence the download latency or the content retrieval time of the peers in the network. Further, 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 as 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.

Although, apart from the topological issues, there are certain other important issues like, service discovery, indexing and replication, and security issues that are also being considered by the research community, however, the focus of my research is centered on improving performance like search, bandwidth wastage and download latency in heterogeneous p2p networks through better topology management schemes.

In our approach, we initially attempt to gain insight into the existing super-peer based network topologies (Gnutella, Kazaa etc.) and also Bit Torrent based p2p systems, 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 attempt to propose suitable topology optimization mechanisms and validate their improvement in performance using simulations.

The work done in the direction of achieving the objectives can be summarized as follows:

1. We have initially studied the impact of network topology on the traffic redundancy generated at the peers in super-peer based networks like Gnutella. We have modeled the effect of the topology formed in Gnutella-like networks on the redundancy in the system. Redundant traffic wastes precious bandwidth, and hence modeling the relation between topology and traffic redundancy helps us to optimize topology so as to reduce bandwidth wastage due to redundancy.

2. We have also studied the impact of topology onthe coverage of the peers in unstructured p2p networks that use flooding as the underlying search mechanism. We derived the coverage of the peers in case of Gnutella-like networks, when the peers use a TTL-based flooding. Our analysis reveals that the coverage of the peers is reduced by the presence of certain edges, that we term as cross and back edges, which are naturally formed when the topology evolves. We derived the probability of occurrence of these edges for various kinds of random networks. The results indicate that the probability of occurrence of these edges increase enormously with increasing distance from the source nodes, which can result in huge traffic redundancy, thus questioning the effectiveness of larger TTL based search.

3. Based on these findings, we proposed 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.

4. For Bit Torrent based p2p networks, we have established the importance of topology in determining the network parameters like download latency and fairness — contrary to the previous belief that topology does not play much role in shaping the download performance of the peers. Based on these analyses, we have been able to derive the topological characteristics of Bit Torrent based systems that improves the average download latency of the system by around 13-17%, whereby improving the fairness of the system.

Kamalesh Ghosh
Email: kghosh@cse.iitkgp.ernet.in
Joined the department in: December 2007

***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 later 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.*

Kunal Banerjee
Email: kunalb@cse.iitkgp.ernet.in
Joined the department in: July 2009

***Kunal Banerjee** received a B.Tech. degree in Computer Science & Engineering from Heritage Institute of Technology, Kolkata in 2008. From November 2008 till July 2009, he worked in Tata Consultancy Services Ltd., Kolkata, as an Assistant Software Engineer. Since July 2009, he has been a research scholar in the department of Computer Science & Engineering in IIT Kharagpur. His research interests are in the areas of Formal Verification and Embedded Systems.*

Supervisors: Prof. Chittaranjan Mandal and Prof. Dipankar Sarkar

Equivalence Checking of Parallel Behaviours in Embedded Systems

Embedded systems are prevalent in all modern gadgets ranging from simple portable devices to sophisticated multiboard systems. As a consequence of the increasing complexity of embedded systems design to make them dependable, robust and time-critical, a lot of study has been invested to guarantee their behaviours. Testing fails to reach this goal due to the exponential number of input cases, which has led to an emphasis on formal verification. The challenges posed by high-performance and large-scale computing have shifted the focus of modern designers from serial to parallel paradigm. Given some behavioural specification, our objective is to verify a system which realizes it. Since we are concentrating on distributed systems containing multiple autonomous computers that communicate through a computer network, or a system with multiple processors that communicate over a bus, there is a need to find a suitable distributed model of computation to model such systems. Our initial choice was *'Multiway Decision Graph'* which has found applications in formal verification. *'Kahn Process Networks'* (KPN) which exhibits deterministic behaviour that does not depend on the various computation or communication delays, seemed to be a better choice and hence all our current studies entail this model. There has been some work where a behaviour, described by a KPN, is realized on some hardware, say a network-on-chip or Cell BE multi-processor. A run-time system, named *'Nornir'*, has been developed (not by us) for executing KPNs on multi-core machines. This system is insubstantial because it is fully software-based and does not concern with any hardware issues such as those that may arise due to computers connected through a network. Hence there is a need to construct a tool that can alleviate these problems.

Mahesh Raghunath Shirole.
Email: mr.shirole@cse.iitkgp.ernet.in, mrshirole@vjti.org.in
Joined the department in: July 2010

Mahesh Shirole *received a B.E. degree in Computer Science and Engineering from Walchand College of Engineering, Sangali in 1996 and an M.E. degree in Computer Science from Veermata Jijabai Technological Institute, Mumbai in 2004. From August 1998 till April 2000, he worked in Bharati Vidyapeet's College of Engineering, Navi Mumbai, as a Lecturer. Since April 2000, he has been in Veermata Jijabi Technological Institute, Mumbai, as a Lecturer. Since July 2010, he has been a research scholar under Quality Improvement Program (QIP) in the department of Computer Science & Engineering in IIT Kharagpur. His research interests are in the areas of Software Engineering, Object Oriented Technology, and Evolutionary Algorithms.*

Supervisor: Prof. Rajeev Kumar

Model Based Test Data Generation using Evolutionary Algorithm

Software testing is expensive, typically consuming roughly half of the total costs involved in software development. Software testing cannot ensure software is bug-free, but it can increase software quality. Generally, the goal of software testing is to design a set of minimal number of test cases such that it reveals as many faults as possible. An automated software testing can significantly reduce the cost of developing software. Software testing is one of the significant component of software engineering with many complex and interrelated constraints. Efficient testing requires systematic and automatic test data generation. Automation comes with cost. It solves the manual test execution problem but creates a new test automation production problem in terms of test data generation. Automating the test data generation process is vital to advance the state-of-the-art in software testing. For years, many researchers have proposed different methods to generate test data automatically, i.e. different methods for developing test data/case generators. Yet it remains an open challenge.

Search-based software testing is an approach to apply evolutionary algorithms i.e. metaheuristic search techniques like genetic algorithms, simulated annealing and tabu search to software testing problems. It is inspired by the observation that many activities in software testing can be formulated as optimization problems. Due to the computational complexity of these problems, exact optimization techniques of operations research like linear programming or dynamic programming are mostly impractical for large scale software testing problems. Because of this, researchers and practitioners have used metaheuristic search techniques to find near optimal or good-enough solutions. Test-input generation often occurs when an implementation of the program under test is available. However, before a program implementation is available, test inputs can also be generated using specification documents or analysis and design models. Commonly, searching for an input in a pool (domain/set) of possible input data is dealt with as an optimization problem.

Model based testing is the development of testing artifacts on the basis of UML models. Source can be seen as a concrete representation of a system and UML models are more abstract representations of the same system. Code-based testing is concerned with identifying test scenarios that satisfy given code coverage criteria, and exactly the same concepts can be applied to more abstract representations of that code, i.e., the UML models. The entirety of all UML models specifies the system completely and sufficiently. An integrated framework is required from UML model to derive test scenarios and test data.

The focus of our research work is that of employing evolutionary algorithms for generating and evolving test cases/test data from UML models. The main goal is to develop framework of automated test case generation using genetic algorithm for UML models. Figure 1 shows the framework for software test data generation.

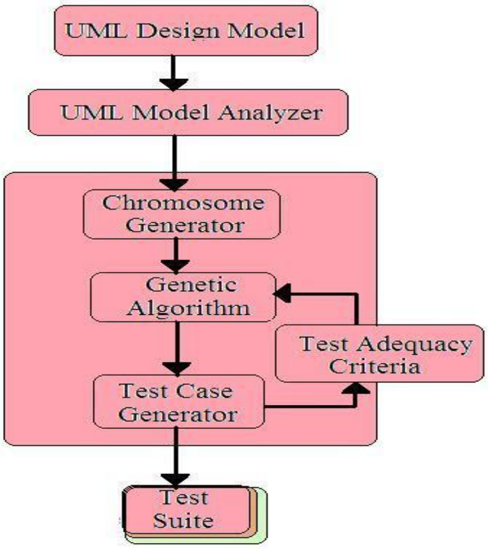


Figure 1: Framework for software test data generation**.**

References

1. P. McMinn, "Search-based Test Data Generation: A Survey". Journal on Software Testing, Verification and Reliability, Volume 14, Issue 2, pp.105-156, 2004.

2. P. Tonella, "Evolutionary Testing of Classes". International Symposium on Software Testing and Analysis, pp.119- 128, 2004.

3. S. Wappler, J. Wegener, "Evolutionary Unit Testing of Object-Oriented Software Using Strongly Typed Genetic programming". In: Proceedings of the 8th annual Conference on Genetic and Evolutionary Computation, pp.1925-1932, ACM New York, (2006).

MANJIRA SINHA
Email: manjira@cse.iitkgp.ernet.in
Joined the department in: July 2009.

Manjira Sinha *received B.Tech. degree in Computer Science from Heritage Institute of Technology, Kolkata in 2009. Since July 2009, she has been a research scholar in the department of Computer Science & Engineering in IIT Kharagpur. Her research interests are in the areas of Cognitive Science and Human-Computer Interaction.*

Supervisor: Prof. Anupam Basu

Human-Computer Interaction and Cognitive Science

From the beginning of the computer era, much research has been done to make the computer more advance an efficient, as a result the capability of computers in terms of complex functions; huge storage and fast processing speed are increasing. From the 20th century, the human factor was started to be considered. The need to effectively 'interact' with computers in order to address the design problems which involves strong sensory-motor features led to the subject Human-Computer Interaction (HCI). A good User Interface can be a way to interact. The term 'good' can be defined in many ways depending on the requirement and the context as mentioned below.

The human side can be modeled with the help of cognitive science. The multidisciplinary subject cognitive science is dedicated towards searching the answers of some questions which have long bothered the human race through ages - 'what's going on there in our mind?' or 'what are the basic simple parameters using which the human brain has developed it's complex formulas?', 'how to represent the cognitive functions which are so similar yet whose variations make every human unique?' ,'what is the very thing consciousness?', 'mere behaviorism or the embodied cognition?' or simply how a person recognize the color red seeing a rose!

Cognitive load associated with navigating in a user interface is an overlapping region of the above two areas. Cognitive Load can be defined as the pressure on the working memory (human ram) while doing an executive control or the amount of information that can be retained in the short-term memory (human cache)**.** It can be of two types: 1) Intrinsic: the inherent load of that process and 2) Extrinsic: the load which can be optimized by an interface design. GUI are mainly of two types: 1. WIMP (Windows, Icon, Menu, Pointing Device) i.e., the familiar desktop environment and 2. Web interface.

Among existing approaches of UI design and evolution often the cognitive capabilities or limitations of the user have not been addressed through a formal framework. Many times it happens that due to constraints present in the cognitive system a user fails to take the advantage of a UI with rich feature sets. An integrated modeling approach is therefore necessary which will formally address both the users' cognitive model and model of the UI. One way can be to model them as two interacting processes or systems. Now, an UI has many parameters on which optimization can be done – Navigability represents the 'ease' with which the user can arrive at the desired goal state. It can be represented in the form of a directed graph network with different states as vertices and edges between two vertices are defined if there exists a way to reach one from another.

The challenge is whether it is possible to come up with some kind of metric or formal system to model the effect of design of a user interface on the cognitive load encountered by that person.

Another area to study can be regarding the persons who are differently able (like mentally challenged or physically challenged) than a 'standard' user. 'How the design parameters to be changed – should they be extended from the existing system or redefined completely?' For example, what problem(s) a person with Dyslexia will feel navigating an interface marked 'easy' by a non-dyslexic person or what will be the modeling approach regarding an interface for blind persons, as, for the others the visual stimuli dominates. The cognitive architecture of a human being can be modeled with the help of already existing systems like ACT-R, EPIC, and RACE etc.

The 'goodness' of a user interface must be defined in these above contexts. The target range of users should be taken into account while designing and each term relevant with the interface design should be described according to the capability and need of the specific group.

Currently my research interest is to bring both the user side and the interface side under an integrated modeling approach.

Maunendra Sankar Desarkar
Email: maunendra@cse.iitkgp.ernet.in
Joined the department in: July 2008

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

Supervisor: Prof. Sudeshna Sarkar

Aggregation of User Preferences for Recommender Systems

Recommender systems suggest items to the users. The items may be of different types such as electronic products, movies, music, books etc. Often different users have different tastes these product types. While recommending items to a user, it becomes important to understand the taste of the user and suggest items according to his/her taste. The systems need some input from the users to understand the users' likings for different items. Capturing a user's interest for different product types or subtypes is possible if the user spends significant amount of time with the system. On the other hand, the user will not feel interested in interacting with the system if he/she does not like the recommendations generated by the system. Quality recommendations are required for both the system and user perspective.

I am focusing on developing algorithms for generating personalized recommendations for users. There are recommender systems which allow users to 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. Netflix, Movielens allow users to rate movies, Amazon allows users to provide ratings for items they have purchased. Similarly there is Jester for jokes, BookCrossing for books, several news portals for news articles etc. The rating assigned by a user to an item may be treated as the utility of the item for the user. If the system can predict the (personalized) utilities of different items for a user, then that information can be used for recommending items to that user. This argument has given rise to the *rating prediction problem* in recommender systems where the task is to predict the rating that a user would give to an item that he/she has not rated in the past.

Collaborative filtering is a widely used technique for rating prediction in recommender systems. Based on the assumption that users with similar tastes would 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. User-based versions of memory based collaborative filtering algorithms assign weights to the users to capture similarities between them. The weighted average of similar users' ratings for the test item is output as prediction. Item-based methods work analogously; the final prediction is a weighted average of the ratings assigned to the different items by the target user. For user-based approaches, the weight of a user is a measure of confidence on that user's opinion. The confidence can be biased if the number of available ratings is small. Rating of an item is dictated only by the ratings given to that item by the similar users. Moreover, if not many users have rated the test item, then the estimate can be poor as it is the weighted average of a small number of ratings. The same problem is shared by item-based approaches as well. Data sparsity is a known challenge for any collaborative recommendation system.

It is understood that collaborative systems looking purely at ratings only have some drawbacks. Even similar users tend to rate the same items differently. This is because rating is subjective, and users have different levels of leniency while rating items. We believe that use of preference relations between items can reduce the bias due to the users' rating habit and may give a clearer understanding about the qualities of the items. Moreover, it might be difficult to pick a particular rating for an item, whereas given two different items it is probably easier to say which one is better (or both are equally good). Keeping these points in view, we propose a memory-based collaborative filtering approach that uses preference relations between items instead of absolute ratings.

In our approach, each user's ratings are viewed as a preference graph. Similarity weights are learned using an iterative method motivated by online learning. These weights are used to create an aggregate preference graph. Ratings are inferred to maximally agree with this aggregate graph. The use of preference relations allows the rating of an item to be influenced by other items, which is not the case in the weighted-average approaches of the existing techniques. This is very effective when the data is sparse, especially for the items rated by few users. Our experiments show that our method outperforms other methods in the sparse regions. However, for dense regions, sometimes our results are comparable to the competing approaches, and sometimes worse.

Going forward, we would like to modify our method to achieve better prediction accuracy for the dense regions. It would be interesting to see if given a dataset it is possible to automatically determine the ranges of data where preference relation based techniques work well and the regions where absolute rating based methods work better. Another important research direction is to augment predicted utility with some other aspects such as novelty and freshness to come up with the final recommendation list.

Parantapa Bhattacharya
Email: parantapa@cse.iitkgp.ernet.in
Joined the department in: July 2010

***Parantapa Bhattacharya** received his B.E. degree in Information Technology from Bengal Engineering and Science University, Shibpur in 2008 and M.Tech. degree in Information Technology from IIT Kharagpur in 2010. Since July 2010 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 Online Social Networks, Cloud Computing, and Network Security.*

Supervisor: Prof. Niloy Ganguly and Prof. S. K. Ghosh

High Performance and Economic Computing Frameworks for Rented Clouds

Companies often face the problem of under or over provisioning of computing resources for their upcoming sites. As the site gets more and more popular it gets increasingly difficult to scale up and meet the demands without exhausting the company's economic resources. This problem is especially acute for social network sites which can get popular overnight. Moreover the load on servers does not remain constant and varies with time and location of request origins. Rented Clouds provide a method for dynamic allocation of computing resources but are quite costly if not used judicially.

It is thus necessary to develop frameworks for economically judicious use of resources from the cloud while not degrading end user experience. The problem is to develop load balancing and optimizing techniques (optimizing computation costs, storage costs, bandwidth costs, and response time) for servers on the cloud.

To maintain a high performance response time it is natural to replicate the servers and perform load balancing. But this technique incurs high cost of storage and inters server bandwidth requirements. To reduce these costs one can perform partitioning of the total data on different servers. But then the servers become heterogeneous and load balancing becomes a difficult issue if one server gets overloaded. The challenge lies in load balancing in a heterogeneous server environment i.e. every server does not have the same data. The current focus is on developing a dynamic system which employs partitioning, replication, and load balancing simultaneously. Social networks sites have an additional advantage to leverage the natural grouping and social relation of their end users to optimize systems even further.

Prasenjit Mondal
Email: prasenjitm@cse.iitkgp.ernet.in
Joined the department in: January 2009

Prasenjit Mondal *received a M.Sc. degree in Computer Science from Vidyasagar University, Midnapore in 2005, and an M.Tech. degree in Computer Science and Engineering from Haldia Institute of Technology, Haldia in 2008. From July 2008 till December 2008, he worked in Telemedicine Project, IIT Kharagpur, as a Junior Project Officer. Since January 2009, he has been a research scholar in the department of Computer Science & Engineering in IIT Kharagpur. His research interests are in the areas of Medical Image Processing.*

Supervisors: Prof. Jayanta Mukhopadhyay and Prof. Shamik Sural

Measuring Lateral Ventricle Width from Neonatal Brain Ultrasound Images

Nowadays, medical image processing has substantially improved the healthcare services by means of computer vision based techniques, for example, automatic analysis of X-ray, ultrasound, CT, MRI images are widely used in computer based processing. An automatic scheme to measure the width of lateral ventricles from the ultrasound images of neonatal brain in coronal view can be helpful.

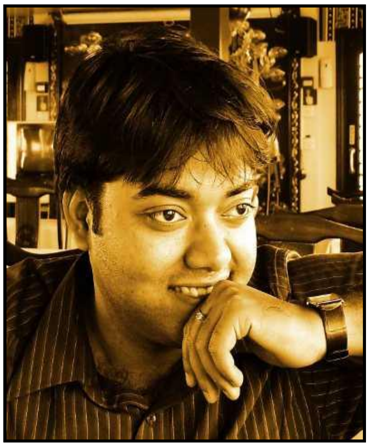
Fetal Ventriculomegaly is a commonly detected abnormality of the fetal brain, and is detected in around 1-2 per 1000 fetus [1]. Isolated Ventriculomegaly occurs in about 0.39%-0.8% of births [1]. Recently, real time ultrasound has been widely used in neonatal intensive care units as a noninvasive diagnostic tool to diagnose intraventricular haemorrhage (IVH), hydrocephalus and to measure intracranial ventricular size. Ventricular size has been measured in neonates in an effort to diagnose early fetal hydrocephalus.

Generally, radiologists use to measure the lateral ventricle width from ultrasound images with the help of USG machine. But other than radiologists and experts it is very difficult for others to measure the lateral ventricle width from the original ultrasound images.

Ventriculomegaly is one of the most sensitive markers for abnormal development of the fetal central nervous system. So assessment of the size of lateral ventricles has become an important part of routine neurological assessment in neonates. We have proposed a method for measuring the width of lateral ventricles in term and preterm neonates from the coronal view of their brain ultrasound images. Based on the ultrasound image, the following steps are performed. 1) Ultrasound image is preprocessed to reduce the speckle noise. 2) The region of interest is extracted from the original image. 3) The image containing lateral ventricles and cavum septum pellucidum (CSP) is segmented. 4) Lateral ventricles and CSP are labeled. 5) Morphological operations are done to smooth the boundaries of lateral ventricles and to fill the cavities. 6) Lateral ventricle width is measured by means of the proposed image processing method.

References

1. J. M. Rennie, C. F. Hagmann and N. J. Robertson, “Neonatal Cerebral Investigation”, Cambridge University Press, 2008.



Priyankar Ghosh
Email: priyankar@cse.iitkgp.ernet.in
Joined the department in: March 2007

***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. Partha Pratim 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.

1. **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.

2. **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.

3. **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.

Rajiv Ranjan Suman
E-mail: rrsuman2001@gmail.com
Joining month and year in the department: July 2007

Rajiv Ranjan Suman *received his B.E. degree in Computer Science & Engineering from Birsa Institute of Technology (BIT), Sindri, Dhanbad (Jharkhand) in 1991, and M.Tech degree in Computer and Information Technology from IIT Kharagpur in 2002. From Mar 1992 to Dec 1995, he worked at BIT Sindri as a part-time Lecturer. He joined as a Lecturer at NIT Jamshedpur (Jharkhand) in Jan 1996. At present, he is an Assistant Professor in the CSE department of the same institute. In July 2007, he joined the Department of CSE, 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 Models of Software Components

Component-based software development derives benefits from software reuse. Large software is built by integrating pre-fabricated, independently developed and reusable executable units, called software components that are usually available as commercial off the shelf (COTS) components. Such components are usually integrated to applications through an application program interface (API). Only a brief description of the functionality and usage syntax of a component, usually in the form of interface description language (IDL) specifications, 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.

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 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 replacement. State-based bugs are difficult to detect using traditional testing techniques [1]. 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 [2]. 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, FSMs lack hierarchy and concurrency and suffer from state explosion problem. Statecharts are compact, expressive, compositional and modular [3]. They naturally tackle the problems and limitations 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. Such conversion can also help in enhancing the understandability of model behavior, automatic generation of code [4] and test cases, etc. 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 two efficient methods and their implementations for converting FSM models to statechart models. The first one is based on analysis of transitions having identical labels. The second one is based on the path analysis of the state transition diagram of the FSM. Our methods detect the possibility of introducing hierarchy and concurrency in the available FSM or statechart and implement them by construction of super OR-states and AND-sates in the resulting statecharts. We implement tools based on these methods that construct nested super states (if possible) in statecharts making them hierarchical and highly compact.

References:

1. Binder, R.V. “Testing Object-oriented Systems: Models,Patterns, and Tools”, Addison-Wesley Longman Publishing Co, Inc, Boston, MA, 1999.
2. 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.
3. Harel, D. “Statecharts A visual formalism for complex systems”, Science of Computer Programming, 1987, Vol. 8(3), pp. 231-274.
4. Benowitz E., Clark K., Watney G. “Auto-coding UML Statecharts for Flight Software”, n SMCIT'06: Proceedings of the IEEE International Conference on Space Mission Challenges for Information Technology, July 2006, pp. 413-417.

R.Rajendra Prasath
Email: rajendra@cse.iitkgp.ernet.in
Joined in: January 2009

***R.Rajendra Prasath** received M.Sc (Mathematic)] from Ramanujan Institute for Advanced Study in Mathematics, M.Tech (CSE) from Indian Institute of Technology, Kharagpur and Ph.D (Mathematics-Computer Science) from University of Madras. Rajendra started his research career with a guest faculty position at University of Madras in 1998. During 2004-2006, he worked as an Assistant Professor at MNMJEC under Anna University, Chennai. Later Rajendra joined Communication Empowerment Laboratory of IIT Kharagpur as a Senior Project Officer. During August 2009 – September 2010, Rajendra was associated with the Norwegian University of Science and Technology (NTNU), Norway as an ERCIM Alain Bensoussan Fellow. Rajendra was a Visiting Fellow at The Artificial Intelligence Research Institute (IIIA), Spanish National Research Council (CSIC), Barcelona, Spain and Swedish Institute of Computer Science (SICS), Kista, Sweden during May - June 2010. Earlier, Rajendra was a University Research Fellow at University of Madras, from November 2001 to April 2003. He also developed tools for Cross Lingual Information Access system (at IIT KGP) as a part of DIT, Govt. of India sponsored research work. Presently Rajendra is a Technical Editor of the journals: Advances in Information Sciences and Journal of Computer Science. Rajendra served as a reviewer for journals: IEEE/ACM Transactions on Networking, Information Sciences, Journal of Convergence Information Technology and several popular international conferences. He is a member of World Federation on Soft Computing, Information Retrieval Facility – Vienna, International Rough Set Society (IRSS) – Warsaw and Information Retrieval Society of India. His research interests include cross lingual information retrieval, textual case based reasoning, machine learning and distributed algorithms for message passing systems.*

Supervisor: Prof. Sudeshna Sarkar

Knowledge Intensive Approaches in Cross Lingual Information Retrieval

Web is growing with the multilingual content as the users share their knowledge in their own language. Like India and Europe, there are many countries having more than one language and web content is growing in the native language of users. With the growth of vast amount of such multilingual web content, it becomes essential for the users to access the information present across other languages. Research on Cross Lingual Information Retrieval (CLIR) and access have emerged in such scenario to assist the information needs of the different language speaking users who may issue queries in one language by enabling them to access the information written in other languages [1].

A major challenge in Information Retrieval (IR) systems is to deal with incomplete or underspecified information in the form of queries issued by users. The IR systems receiving such queries need to fill in the gaps in the underspecified query of the users so as to improve the retrieval efficiency [2]. The problems multiply in cross lingual domain where CLIR systems often utilize translation to cross the language barriers between a query and the documents. However the query translation and translation disambiguation of underspecified queries make the retrieval and ranking of documents more challenging for cross lingual retrieval. In this scenario, it is important to study more appropriate ranking algorithms for CLIR systems.

Pirkola (2003) proposed the *Query Structuring* method for grouping query keywords. This method further suggests the use of query operators in such a way that more weight is being assigned to important or correct keywords than the other keywords [3]. This query structuring captures the clue on the intention of users' information need with English-Finnish news data. Here we are attempting towards such a mechanism to retrieve documents by identifying the query type that gives a clue on the information need of the users across multiple language contents and then to rank these retrieved documents in a better order. We plan to test the proposed approaches on the standard benchmark IR test collections provided by TREC, CLEF and FIRE and standard TREC evaluation metrics will be used to measure the retrieval efficiency [4].

Initially we restrict our focus only on Indian tourism related web documents. Already we have crawled such documents and applied information extraction using *link-to-text* ratio based heuristic. This method eliminates most of the noisy contents from web documents and extracts text information which is then indexed. Then for a selected set of FIRE queries (in English), we performed the retrieval of web documents. We have, in hand, the relevant judgment for a small set of 1200 web documents [manually judged]. The retrieval accuracy, after and before noise filtering, has been measured with Mean Averaged Precision (MAP) and Precision@top *d* documents (P@d) using TF-IDF and BM25 algorithms. The preliminary results are promising with monolingual web collection. We further plan to perform this experiment on cross lingual web collection and to measure the effectiveness of ranking after noisy text removal.

Recently we have developed distributed representations to detect higher order term correlations in textual content. This method efficiently indexes words and its context in the form of a feature vector whose length is initially fixed up. Then similarity of documents with the given query is computed based on the similarity between their constituting features in the query and the documents. Experimental results show that applying distributed representations captures higher order term correlations efficiently. Next we plan to apply this method with possibly variations on three text datasets from FIRE collections, preferably in English, Bengali and Tamil languages.

References:

5. P. Majumder, M. Mitra, and K. Datta, "Multilingual information access: an Indian language perspective", in: Proc. ACM SIGIR Workshop on New Directions in Multilingual Information Access, Seattle, 2006.
6. D. He, D. Wu, "Enhancing query translation with relevance feedback in translingual information retrieval", Information Processing and Management, Vol. 47, 2010
7. A. Pirkola, D. Puolamaki, K. Jarvelin, Applying query structuring in cross-language retrieval, Information Processing and Management, Vo. 39, 2003.
8. M. Sanderson, "Test collection based evaluation of information retrieval systems", Foundations and Trends in Information Retrieval, Vol. 4, No. 4, 2010.

Rajib Ranjan Maiti
Email: rajib.maiti@gmail.com
Joined the department in: July 2009

Rajib Ranjan Maiti*receiived a M.C.A degree from Biju Patnayek University of Technology, Orissa in 2004, and an M.Tech. degree in Computer Science from Indian Institute of Technology, Kharagpur in 2008. From July 2008 till June 2009, he worked in Magma Design Automation India (Pvt.) Limited, Bangalore, as an Associate Member of Technical Stuff. Since July 2009, he has been a research scholar in the department of Computer Science & Engineering in IIT Kharagpur. His research interests are in the areas of Mobile Networks.*

Supervisors: Prof. Niloy Ganguly and Prof. Arobinda Gupta

Analyzing the Performance of Epidemic Routing in DTNs on Human Mobility

The recent proliferation of wireless devices has opened up new paradigms of networked systems. One such type of network that has seen significant interest from researchers is Delay Tolerant Networking (DTN). A DTN consists of mobile wireless devices that occasionally come into contact with each other. The primary characteristic of such networks is that no end-to-end path may exist between a pair of agents at any particular instance of time. Communication in such networks follows the store-carry-forward paradigm, where an agent stores a packet temporarily if it cannot forward it towards the destination; the packet is forwarded to when it comes in contact with another agent following some routing policy.

Routing in a DTN can typically be achieved using some variation of the epidemic protocol. In the classical SIRS model of epidemiology that we address in this work, a recovered agent moves back to the susceptible state after a certain time. A susceptible agent can only switch to the infected state on reception of a message from a neighbor. In the recovered state, the agent is idle in the sense that it neither transmits nor receives any message even if there are agents in its contact range or it is in the contact range of other agents. After the idle period the agent becomes susceptible again. Hence the model is known as the SIRS model of epidemics. Among the different mathematical models of epidemics, the SIRS model is better suited to DTN communication because of the recovered state, allowing energy-strapped mobile agents to conserve some of their energy. The time management for each state is crucial to ensure that the message stays in the network until the recipient of the message gets it and the energy is used efficiently as well.

Routing in a DTN has earlier been studied widely using omnidirectional antenna (OA).These reported works using OA can broadly be classified into two categories. A number of approaches have been taken to reduce the overhead and improve the performance of epidemic routing by implementing probabilistic forwarding, redundancy suppression scheme as well as schemes which utilize history. There have also been several works which, mainly through extensive simulation, have shown the performance of routing under different choices of parameters (e.g. message holding period, transmission range, etc). Moreover, the mobility model used was only the Random Walk Mobility (RWM) model, which does not represent any real world movement pattern. Hence the results, though illustrative, are not indicative of the expected performance of epidemic protocol with directional antennas (DA) in real world scenarios.

In this work, we analyze the routing as well as broadcasting performance of the epidemic protocol for realistic mobility models in a DTN setup with both OAs and DAs. In particular, we study how the average message delivery delay and the average number of hops to reach a destination of the epidemic protocol vary with distance from the sender. The realistic mobility model considered is the real human daily life movement patter which some researcher have termed as Self Similar Least Action Walk (SLAW) [1] model of mobility. SLAW has been derived from daily lifehuman mobility patterns and hence is representative of a practical, real world mobility pattern.

The primary contribution of this work is to show that the use ofdirectional antenna gives better performance in terms of routing andbroadcasting than using only omnidirectional antenna in realisticmobility models. The effects of different parameters such as agentdensity, range of OA and DA, antenna rotational probability and thefraction of agents using DA, on the performance of the system are alsostudied and analyzed. The interesting observation in case when all agents have OA is that the self similarity parameter (i.e. Hurst parameter) of human walk is proportionally related with both themessage delivery delay and the number of message handovers for any fixed transmission range. The observation in case of a system where a fraction of agents are using DA is that when system size is larger butthe agent density is low, DA performs much better in terms of deliverydelay and number of message handover than that of the system where allthe agents have OA.

References

1. Kyunghan Lee, Seongik Hong, Seong J. Kim, Injong Rhee and Song Chong, “SLAW: A Mobility Model for Human Walks”, Proceedings of the 28th Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM), Rio de Janeiro, Brazil April, 2009.

Rishiraj Saha Roy
Email: rishiraj@cse.iitkgp.ernet.in
Joined the department in: December 2009

***Rishiraj Saha Roy** received a B.E. degree in Information Technology from Jadavpur University, Kolkata in 2007, and an M.Tech. degree in Information Technology from Indian Institute of Technology Roorkee, Roorkee in 2009. Since December 2009, he has been a research scholar in the department of Computer Science & Engineering in IIT Kharagpur. His research interests are in the areas of Information Retrieval and Complex Networks.*

Supervisor: Prof. Niloy Ganguly

Understanding and Leveraging the Structure of Web Search Queries

Current search engines consider a query to be a bag-of-words and assume that a relevant document will have all or most of the keywords; stop words such as *in*, *of*, and *why*, are ignored altogether. Motivation for this work stems from the fact that there is much more inside a query than just its constituent terms. For example, the query "*can't view large text files in windows 7*" is definitely not the same as an unordered list of its constituent terms – *can't, view, large, text, files, windows* and *7*. More often than not search engines return very unsatisfactory results for this type of queries, because they ignore the facts that here *large text files* is an entity, *can't view* is an action on *large text files*, and *in* indicates that the rest of the query is in the context of the *windows 7* operating system. Ironically, this seems to happen when the user tries to specify the information need a bit more precisely.

The aim of the proposed research is to understand the underlying structure of queries, learn those structural units and patterns automatically from data and apply this knowledge to improve the performance of search engines. We can intuitively feel that English language grammar, which is so essential to the understanding of natural language phrases and sentences, does not hold for Web queries. If we compare a Web search query and its corresponding natural language sentence or phrase, we would often observe that many words have been dropped while forming the query. Also, there is more flexibility in the relative ordering of the words in the query – two queries can be semantically equivalent even if the ordering of the words varies to a large extent. These issues propel us to formulate a new grammar for queries – which we can extract from the data, based upon our structural organization. Once we have a working definition of a grammar in place, the next task would be parsing a query in accordance with this grammar. We foresee that if we are able to grasp the internals of a query at this level, we can use this knowledge to bring about great improvements in search quality by enhancing established techniques like query expansion and re-ranking the list of search results.

To this end, we plan to apply machine learning techniques on data resources such as query logs, click-through data, per user sessions' data, and the contents of Web documents. This would require rigorous manual analysis of query logs to understand the structural patterns of queries. Past work has talked about intent of a query as a whole [1, 2]. But our initial study shows us that the words in a query itself can be grouped into two classes, which we shall call *content* and *intent* words (or phrases). While content words are like keywords that must be matched at the document side, intent words can be used to guide the search engine in other ways. We note that labelling as content or intent is not at the word level but for meaningful expressions as a whole. This motivates us to devise a suitable scheme to perform query segmentation (breaking a query into its meaningful segments). After observing and annotating a large amount of query logs, we came up with a robust linguistic classification of intent words. These rules were derived from first principles and based on the nature of interaction between content and intent words. We found that well-established statistical techniques can be used to perform query segmentation (with the segments thus obtained aligning satisfactorily with our notions of content and intent) as well as distinguish between content and intent words. Our results also have a good degree of concordance with data annotated by humans.

We propose to make significant progress in the foregoing lines of thought. We believe that the overall idea is capable of introducing a new paradigm in Web search – trying to understand the meaning of a user query from its structure before actually diving in to retrieve the results. We also foresee that as we go along, we would also be able to shed light on various other interesting phenomena – the learning curve of users when it comes to being successful in Web search, search patterns of users from different geographical locations, and customizing results based on search patterns of individual users.

References

1. Broder, "A Taxonomy of Web Search", ACM (Association for Computing Machinery) SIGIR (Special Interest Group on Information Retrieval) Forum, Volume 36, Issue 2 (Fall 2002), 2002, ACM, New York, USA, pages 3 - 10.

2. J. Jansen, D. L. Booth, and A. Spink, "Determining the informational, navigational, and transactional intent of Web queries", in Information Processing and Management (IPM), Volume 44, Number 3, May 2008, Pergamon Press, Inc., New York, USA, pages 1251 - 1266.

Ruchira Naskar
Email: ruchira@cse.iitkgp.ernet.in
Joined the department in: July 2010

***Ruchira Naskar** received a B.Tech degree in Information Technology from West Bengal University of Technology in 2008, and an M.Tech degree in Information Technology from Indian Institute of Technology, Kharagpur in 2010. Since July 2010, she has been a research scholar in the department of Computer Science & Engineering in IIT Kharagpur. Her research interest is in the area of reversible digital watermarking.*

Supervisor: Prof. Rajat Subhra Chakraborty

Reversible Watermarking of Digital Images

Digital watermarking is the act of hiding some information into any kind of multimedia data. This technique has been long used to protect the intellectual property rights of images. Some information (termed as watermark) is embedded into a cover image, in such a way, that distortion of the cover image due to watermarking is almost negligible perceptually.

Over the last couple of decades, lots of researchers have focused their interest on *reversible watermarking*. These techniques allow the image restored after the watermark extraction, to be same as the original cover image, pixel by pixel, bit by bit. Reversible watermarking is of utmost importance in military and medical imagery, where recovery of the original image after watermark extraction is of as much importance, as proving authenticity of the image. But in many cases, in spite of using reversible watermarking technique, bit by bit recovery of the cover image becomes impossible. For example, in military applications, images sent as data packets or frames over highly noisy channels, exhibit distortions after watermark extraction. The channels, over which military data are sent, are generally highly noisy. Researches indicate that packet error rate of such channels can be as high as 30%.

In the preliminary parts of my research, I have investigated, the amount of distortion of a grayscale image, when it is reversibly watermarked, then sent over such a highly noisy channel, and finally recovered at the receiver side. We have implemented the main classes of reversible watermarking schemes, and compared the distortion levels produced, in the above scenario. Interestingly, we found that the amount of distortion induced by one such class of algorithms is even lesser than the distortion experienced by an image without any watermarking. My research is going to probe into the theoretical analysis of the above effects, as well as on reversible watermarking of colored images.

Sabyasachi Karati
Email: sabyasachi.karati@gmail.com, skarati@cse.iitkgp.ernet.in
Joined the department in: June 2010

***Sabyasachi Karati** received his B.E. degree in Computer Science & Engineering from Jadavpur University, Kolkata in 2008 and M.Tech. degree in Computer Science & Engineering from IIT Kharagpur, Kharagpur, West Bengal in 2010. He joined as PhD scholar in the department of Computer Science & Engineering in IIT Kharagpur in June 2010. His research interests lie in the areas of Algorithms, Cryptography and Computational Number Theory.*

Supervisor: Prof. Abhijit Das.

Algorithm Design and Implementation Issues in Cryptography

Currently, we are working on an old problem of *Signature Schemes* in Public-Key Cryptography. We are trying to verify multiple *Digital Signatures* in batches, especially *Elliptic Curve Digital Signatures*. We proposed an algorithm which is based upon the naive idea of taking square roots in the underlying field. We proposed two new algorithms which replace square-root computations by symbolic manipulations to improve the efficiency. We did experiments on NIST prime curves to measure the speedups. We obtained a maximum speedup of above *six* over individual verification if all the signatures in the batch belong to the same signer and a maximum speedup of about *two* if the signatures in the batch belong to different signers, both achieved by a fast variant of our second symbolic-manipulation algorithm. These algorithms are practical only for small (≤ 8) batch sizes. We also port our algorithms to the NIST Koblitz curves defined over fields of characteristic 2.

Sandip Karmakar
Email: sandip1kk@gmail.com
Joined the department in: Jul 2008

***Sandip Karmakar** received a B.E. degree in Computer Science and Technology from Bengal Engineering and Science University, Howrah, WB in 2004, and an MS (By Research). degree in Computer Science from Indian Institute of Technology, Kharagpur in 2010. From July 2004 till August 2006, he worked in TCG Software Services Pvt. Ltd., Kolkata, as a Software Engineer and from November 2006 till December 2007, he worked in Tata Consultancy Services., Kolkata, as a Systems Engineer. Since July 2008, he has been a research scholar in the department of Computer Science & Engineering in IIT Kharagpur. After completing his MS (by Research) in Cryptography under Prof. Dipanwita Roy Chowdhury and Dr. Debdeep Mukhopadhyay, he joined as a PhD scholar in the same department under Prof. Dipanwita Roy Chowdhury. His research interests are in the areas of Cryptography and Network Security.*

Supervisor: Prof. Dipanwita Roy Chowdhury

Design and Analysis of Stream Ciphers

Cryptography pertains in the domain of secure encryptions and decryptions for information communication over insecure channel. Cryptography is generally studied in two areas, symmetric and asymmetric cryptography. Symmetric cryptography is researched in two divisions as well, namely, *block ciphers* and *stream ciphers*. Stream ciphers are encryption algorithms that work best on resource constraint environments. After the standardization of block cipher encryption scheme AES, an effort is going on throughout the world to standardize stream ciphers (*NESSIE[1], eStream[2]*). Another important domain of research in cryptography is analyzing a proposed cipher, called *cryptanalysis*. After a cipher is proposed, generally, it needs to be studied under cryptanalytic techniques till it establishes itself as a potential algorithm. *The area of interest of the present researcher is in the area of design of stream ciphers as well as cryptanalysis of stream ciphers.*

Cellular Automata was established as a good building block for pseudorandom sequence generator [3] or stream cipher. However, mostly linearity or biased correlation of nonlinear rules has made most of the proposed cellular automata based stream ciphers susceptible to cryptanalysis. Our work is in the area of hybrid cellular automata which combines both linear and nonlinear cellular automaton to produce cryptographically robust stream ciphers.

The work also includes the study, design and implementations of other primitives which are expected to be efficient in design of stream ciphers.

Other parts of our work are in the area of cryptanalysis of stream ciphers. The objective is to work on two cryptanalysis techniques, scan based side channel attacks and fault attack.

Scan chains are design for testability (DFT) feature for testing hardware implementations of algorithms. However, it is seen in literature [4] that this testability can be exploited to perform analysis of cryptographic algorithms. The research work includes study of cryptanalysis techniques on other stream cipher algorithms using scan chains. Countermeasures against such cryptanalysis are also included. The aim is to propose a general algorithm to attack any stream cipher by scan chain based attacks.

Fault attacks exploit induction of faults to an implementation of cryptographic algorithm. It is widely shown in literature [5] that cryptographic implementations can be practically broken using fault induction. Most of the fault attacks known till date on stream ciphers are using single bit faults. Our work is both in the area of single and multi-bit fault analysis of stream ciphers and their countermeasures.

References

1. The NESSIE Project: https://www.cosic.esat.kuleuven.be/nessie/.
2. The eStream Project: http://www.ecrypt.eu.org/stream/.
3. S. Wolfram. Random sequence generation by Cellular Automata. In Advances in Applied Mathematics, Volume-7, pages 123–169, 1986.
4. M. Agarwal, S. Karmakar, D. Saha, and D. Mukhopadhayay. Scan-based side channel attack on stream ciphers and its countermeasures. INDOCRYPT, 2008.
5. Jonathan J. Hoch and Adi Shamir, Fault Analysis of Stream Ciphers. CRYPTOGRAPHIC HARDWARE AND EMBEDDED SYSTEMS - CHES 2004. Lecture Notes in Computer Science, 2004, Volume 3156/2004, 1-20.

Sanjay Chatterji
sanjaychatter@gmail.com
Joined the department in: July 2008

***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 Bengal Engineering and Science University (Formarly 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 in 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

Bengali Hindi Machine Translation Issues

Machine translation is a process by which a text from one language (source) is translated to a text in another language (target). We consider a sentence as the basic unit of machine translation. A source language sentence can be translated into a target language sentence in two ways: Rule Based Machine Translation (RBMT) and Statistical Machine Translation (SMT). Standard statistical approach is to build a model based on the bilingual parallel sentences. On the other hand, rule based system is built based on a large number of built-in linguistic rules and dictionary entries.

Hindi is the 4th and Bengali is the 8th most widely spoken languages in the world. Though Bengali and Hindi have syntactic similarity, the translation of the Bengali sentence to the Hindi sentence is not always easy. Some of the issues are copula insertion, word and chunk reordering, particle translation, agreement issue, idiom translation, etc.

We wish to investigate both statistical and rule based approaches to build machine translation systems between Bengali and Hindi, based on the resources that are available. Bengali is a resource poor language. For the rule based system we have prepared the baseline systems like POS tagger, Chunker, Named Entity Recognizer, Parser, Morphological Analyzer, Morphological Generator and so on for Bengali with the help of the resources like 200,000 words POS and chunk tagged Bengali corpus, 150,000 words Named Entity tagged Bengali corpus, a dictionary with 20,000 parallel concepts and the large number of rules like Transfer Grammar Rules, Prunning Rules, Morphological Rules and so on. For the statistical system, we have prepared 13,000 parallel sentences and used them in MOSES, Jushua and SAMT open source systems. The problems of one system can be solved by another. Our main goal here is to find the problems of the rule based and statistical systems and solve them in a hybrid framework.

The problems of the sentences which are not translated by rule based system correctly have many types. As the Bengali Hindi rules are not comprehensive, some problems are coming due to the lack of rule set. Some problems are dependent on context information. For example, a word of one language may be mapped with many words of another language. To choose the appropriate word from many choices, we need to use the context information and therefore use the statistical approach. Most of the recent statistical systems are based on the lattice based decoding. The word lattice that is used in our work is a weighted directed acyclic graph with one start node and one end node. In this proposed statistical approach the word that appear more frequently in a corpus in terms of individual occurrences and contextual information, is selected to be replaced. We have improved the lexical selection of Bengali - Hindi rule based machine translation system with the proposed model.

The outputs of statistical system have 3 kinds of problems. Some are not translated to target language, some are wrongly translated and some which are not to be translated have been translated (names, foreign words etc). We have tried to translate the non-translated words using the dictionary and rules. In this hybrid architecture the phrase alignment table of a baseline Bengali Hindi statistical system is enhanced by the dictionary and a parallel name list. The decoder is executed using the enhanced phrase table and it translates some more words. This output is further processed by applying some affix based postprocessing rules.

There are several other hybrid approaches that need to be explored in the Bengali Hindi translation framework. Similar approaches can be tested for other Indian language pairs and for the pair of the English language and one of the Indian languages.

Santosh Ghosh
Email: santosh@cse.iitkgp.ernet.in
Joined the department in: July 2008

Santosh Ghosh *did his B.Tech from Haldia Institute of Technology (HIT), under the Vidyasagar University, in Computer Science and Engineering, and subsequently joined as a lecturer at CSE department of HIT in 2002. He pursed for further study of Master of Science (M.S.) from the Department of Computer Science and Engineering at the Indian Institute of Technology Kharagpur in 2006. After completion of the Master's course in the year 2008, he joined the same department in the same year, as a PhD student. During this time, his research area has been broadly in the field of Cryptographic Hardware and Side-channel Attacks. His other research interests include VLSI Design and Testing.*

Supervisor: Dr. Debdeep Mukhopadhyay and Prof. Dipanwita Roy Chowdhury

Design and Analysis of Pairing based Cryptographic Hardware for Prime Fields

The primary challenge in modern day cryptographic hardware development lies in coping with progressively strong physical attacks commonly referred to as sidechannel analysis. This research deals with practical implementations and analysis of physical security of pairing based cryptographic operations on prime fields. Pairing computation and elliptic curve scalar multiplication are two major operations in pairing based cryptography. These operations in turn rely on arithmetic in finite fields - prime fields (Fp). Hence, this work first designs a portable and compact architecture for Fp arithmetic. Subsequently, the work proposes an efficient dual-core cryptoprocessor for elliptic curve scalar multiplication based on the above compact Fp core. Field Programmable Gate Array (FPGA) is a relevant platform which provides various in-built features for optimizing arithmetic operations. A configurable core on FPGA device has been developed for Fpk arithmetic based on the above optimized Fp primitive. Two such configurable cores are utilized for developing a pairing cryptoprocessor which computes pairing over Barreto-Naehrig curve. Securities of pairing computations against fault and power attacks are subsequently addressed in this work. The work further studies existing as well as new vulnerabilities of pairing computations against fault and power attacks. Suitable countermeasures are also proposed to resist those attacks.

Saptarshi Ghosh
Email: saptarshi@cse.iitkgp.ernet.in
Joined the department in: July 2009

***Saptarshi Ghosh**received a B.E. in Computer Science from the Bengal Engineering and Science University, Shibpur (erstwhile Bengal Engineering College) in 2005, and an M.Tech. in Computer Science from IIT Kharagpur in 2007. In 2007, he joined the Department of Computer Science and Technology, BESU, Shibpur as a Lecturer. Since July 2009, he has been a research scholar (sponsored) in the department of Computer Science & Engineering, IIT Kharagpur. His research interests are in the area of Complex Network Theory, specifically in Online Social Networks and Transportation Networks.*

Supervisor: Prof. Niloy Ganguly

Effects of a Soft Cut-off on Number of Links in the Twitter Online Social Network

Online Social Networks (OSNs) are now among the most frequently accessed sites on the Web. Apart from communicating with friends and others sharing common interests from all over the world, OSNs are also being used for advertisement, personalized / localized information search and so on. There has been an exponential rise in the number of users and amount of activity in OSNs in the past few years, and the popular OSNs such as Facebook, Twitter, Orkut, Flickr, each have several hundred millions of users at present.

Due to the enormous rise in the number and activity of users in recent years, the popular OSNs are experiencing problems of scalability and increasing spam. Several popular OSNs have adopted a common technique to deal with the above issues: they have imposed limits on the number of friends / social links that a user can have in the network (in graph-theoretic terms, restrictions on the degree of nodes in the social network). Such restrictions reduce strain on the OSN infrastructure caused due to member-to-all-friends communication, and also prevent spammers from contacting large number of users indiscriminately. However, due to the increasing levels of user-activity, more and more legitimate users are also getting affected by such restrictions; hence the restrictions are being frequently criticized by the socially active and popular legitimate users of the OSNs, as an encroachment on their freedom to have more friends. Thus the OSN authorities are facing a trade-off in the design of restrictions.

In view of this trade-off, Twitter [1] - one of the OSNs worst affected by the problems of spam and system-overload - has recently imposed an innovative restriction based on a 'soft cut-off' [2]. In contrast to the fixed (or 'hard') cut-offs in other OSNs, the restriction in Twitter gradually adapts to the requirements of users who achieve a certain level of popularity in the social network, thus attempting to minimize dissatisfaction among popular users. We are investigating the effects of this soft cut-off on various aspects of the Twitter OSN, such as on the network topology, system-load (on the OSN infrastructure), and satisfaction of legitimate users and on the strategies used by spammers.

We have observed that the Twitter restriction has significantly altered the topology of the OSN, and have proposed a Complex Network-based model for restricted growth of networks under different ‘hard’ and ‘soft’ cut-offs (paper at WOSN 2010). Subsequently, we have extended this model to develop a complete analytical framework which OSN authorities can use to design restrictions that suitably balance the two conflicting objectives of reducing system-load and minimizing dissatisfaction among legitimate users (paper accepted at IFIP Networking Conference 2011). We find that Twitter's policy well balances both these objectives, but how effective is the restriction against spammers? Through identification and analysis of 4500 spam-accounts in Twitter, we discover that contemporary spammers adopt intelligent ‘collaborative’ link-formation strategies to successfully counter such restrictions and increase the reach of generated spam, such as forming ‘spam-farms’ by linking with thousands of other spammers as well as targeted legitimate users (poster accepted at WWW 2011).

To the best of our knowledge, this is the first set of work on effects of restrictions in OSNs, as well as the first attempt from a Complex Networks perspective to analyze network-growth in the presence of ‘soft’ cut-offs.

References

1. Twitter, “http://twitter.com”
2. Twitter Help Center, Following Rules and Best Practices, “http://support.twitter.com/forums/10711/entries/68916”

Satya Gautam Vadlamudi
Email: satya@cse.iitkgp.ernet.in
Joined the department in: August 2009

Satya Gautam Vadlamudi *was born in Ghantasala, India, in 1986. He received the B.Tech. (Hons.) degree in Computer science and engineering along with a minor in Mathematics and computing from the Indian Institute of Technology (IIT) Kharagpur, Kharagpur, India, in 2008. From June 2008 to July 2009, he was a Software Engineer with Google India Pvt. Ltd., Bangalore, India. Since August 2009, he has been a Research Scholar with the IIT Kharagpur. His areas of interest Include AI, data mining, design and validation of dependable systems, and algorithm design. Mr. Vadlamudi was the recipient of the Pratibha Award from the Government of Andhra Pradesh (2004), the MCM Scholarship (2004–2008), and the SAP Labs Doctoral Fellowship (2010–present) for his academic and research works.*

Supervisor: Prof. Partha Pratim Chakrabarti

Anytime Search Methods for Solving Optimization and Coverage Problems

Search techniques are widely used for solving diverse computationally hard problems such as combinatorial optimization, data mining, data collection/processing (for example, web), etc. In most of the cases, finding the optimal solution(s) using basic search methods requires exponential order resources (time, memory), often making such (basic) methods inapplicable/non-scalable. For handling large problem instances, it is essential to come up with methods which can work with given memory and produce best possible results in a given time frame.

Algorithms which produce good sub-optimal solutions quickly, and improve upon them when given more time are known as *anytime* algorithms. Algorithms which work with the given memory to produce results are known as *memory-bounded* algorithms. Algorithms with both the aforementioned features are referred to as *memory-bounded anytime algorithms*. Design of good memory-bounded anytime algorithms are vital in many application domains, where finding an optimal solution can potentially takes years of time. A major focus of our work relates to the development of memory-bounded anytime algorithms for a variety of combinatorial search problems.

Optimization problems such as traveling salesman problem, knapsack problem, etc. have exponential order search spaces, and hence finding the optimal solution is often time consuming. The goals in these problems are to find minimum/optimum values of a function and/or the path(s) to reach a goal state from an initial state. Also, many application domains such as design and validation of fault-tolerant embedded systems, spatio-temporal data mining, etc. involve problems with huge search spaces, where anytime heuristic methods can play a vital role. Consider the validation of an embedded control system with respect to errors such as small amounts of shift, noise, and spikes, on different signals in the system (a part of robustness analysis), the number of combinations to be tested in order to validate the controller is very large. Similarly, dynamic reconfiguration of embedded systems involves solving the multi-processor scheduling problem online, which is a well known NP-complete problem. Spatio-temporal data mining has several dimensions to it such as mining valid groups of users, most traversed paths, events, etc., each of which demands for efficient search methods that can scale for large data sets. We develop efficient anytime search methods and apply them to handle the above problems.

We have developed a memory-bounded anytime heuristic search algorithm called MAWA* (Memory-bounded Anytime Window A*) that can be used for solving optimization problems. MAWA* uses the window-bounded anytime-search methodology of AWA* as the basic framework and combines it with the memory-bounded A*-like approach to handle restricted memory situations. Simple and efficient versions of MAWA* targeted for tree search have also been proposed. MAWA* is complete, anytime, and memory bounded. Experimental results of the sliding-tile puzzle problem and the traveling-salesman problem show the significant advantages of the proposed algorithm over existing methods. In future, our work is focused on developing algorithms that can cover the goal-space by producing diverse good quality solutions which would help in, increasing the overall robustness, providing attractive alternatives, etc.

Shyamosree Pal
Email: shyamosree@cse.iitkgp.ernet.in
Joined the department in: July 2008

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 indigital geometry with applications to image analysis and computer graphics.*

Supervisor: Prof. Partha Bhowmick

Estimation of Digital Circularity Based on Number-Theoretic Characterization of Digital Curves

Digital circularity is a well-researched topic for its real-world practicality to circularity measure, estimation of discrete curvature, circular arc segmentation, etc. Several geometric techniques have been proposed over the years to measure digital circularity of an object and to address the related issues. Most of these works, however, had not considered the inherent digital-geometric properties of digital discs/circles, which are indeed difficult to establish. Our work reveals a novel technique to determine whether a digital curve segment is digitally circular using the number-theoretic properties of the digital curves. 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. We have 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 also been shown in our work. Experimental results demonstrate the efficiency and elegance of our proposed technique.

For real-world applications (e.g., vectorization) in which the circular arcs usually deviate from the actual/well-defined digital circular arcs, an approximation algorithm might be more useful. The decision problem on approximate digital circularity is that, given an approximation parameter $\boldsymbol{\tau}$ and a digital curve segment $\boldsymbol{S}$, whether there exists a digital circle $\boldsymbol{\mathcal{C}}^{Z}$(r) for which each point $\boldsymbol{p}$ of $\boldsymbol{S}$ is at most $\boldsymbol{\tau}$ units apart from the corresponding nearest point of $\boldsymbol{\mathcal{C}}^{Z}$(r). A judicious relaxation of the integer interval of radii is required to address this problem. The digital-geometric solution based solely on integer intervals is technically engrossing, which, if achieved, would open up novel possibilities to solve the approximate solution of circular arc segmentation in related applications.

Sk. Subidh Ali
Email: subidh@cse.iitkgp.ernet.in
Joined the department in: January 2009

***Sk Subidh Ali,** is a Phd Scholar in the Department of Computer Science and Engineering, Indian Institute of Technology Kharagpur. His area of research is side-channel cryptanalysis and hardware design security. He has received his Master of Engineering degree in Information Technology from West Bengal Technical University in 2007 and Bachelor of Engineering degree in Computer Science and Engineering in 2003 from Burdwan University. He has served as a lecturer for Bankura Unnayani Institute of Engineering.*

Supervisor: Prof. Debdeep Mukhopadhyay

Design and Analysis of Fault Attack Tolerant Cryptosystem

Hardware implementations of cryptographic algorithms are vulnerable to malicious analysis that exploits the physical properties of the designs. These attacks which exploit the implementation specific weaknesses are known as Side-Channel Attacks (SCA). Information derived from power consumption, electro-magnetic radiation, execution time, and other similar side-channels drastically reduce the complexity of cryptanalysis. Another form of attack which analyzes crypto-devices under accidental or intentional faults to obtain the secret key is known as fault attacks. Fault attacks were first introduced by Boneh et. al., who observed that a single fault in one of the two exponentiations required generating a RSA signature using the Chinese remainder theorem, would allow an attacker to retrieve the private key. Subsequently, Biham and Shamir extended the idea of Differential Fault Analysis (DFA), based on the combination of differential cryptanalysis and fault attack to attack block ciphers, like DES.

These fault based cryptanalysis clearly show the potent threat to the implementation of modern block cipher like AES. There is another form of threat, which lies in association of multiple untrusted parties in design, fabrication and deployment of cryptographic hardware. The nexus between untrusted parties associated in the development can illegally modify the circuit (Trojan) to release the secrets in the form of side-channel leakage. One such example is shown by Lang Li where an inserted Trojan circuitry leaks information through a covert side-channel that lets a team of conspiring malicious parties discover the encryption key.

As with the advent of AES by the NIST, it has become a de facto standard for all the industries in data security. We focus on AES to study the topic of fault attack resistant AES implementations. The present work targets to develop existing fault attacks on AES-128, AES-192 and AES-256 to reduce the number of faults required and the time complexity of the attacks.

The existing fault attack on AES-128 requires a brute-force search of 2^{32} and a time complexity of 2^{32} . We improve this attack by reducing the search space to 2^{8} and reducing the time complexity to 2^{30} . Therefore, our attack on AES-128 is more lethal. We mount a new attack on AES-256 using two faulty ciphertexts. Existing attack on AES-256 requires three faulty ciphertexts.

Most of the recent fault analysis on AES require induction of single byte faults in a specific round of AES. However, there is a high probability that the faults spread to more than one byte. Hence this work investigates the multiple byte faults on AES and analyzes the effect of such faults by classifying the nature of faults. We proposed an improved Multi-byte fault attack on AES-128 which requires brute-force search of 2^{8}, 2^{16}, 2^{24} depending on three basic fault models where as the exiting attack requires brute-force search of 2^{32}, 2^{64}, 2^{96} .

The work investigates the application of fault attacks on multi level attacks in the design flow and the design of hardware Trojans. Trojans are stealthy circuits which leak information to the implanter, who can trigger the Trojan. However to a normal user the Trojan should not be detectable. The simplicity of fault attacks motivates to study the application of fault attacks in the design of Trojans. We designed and implemented a hardware Trojan based on fault attack which takes minimal power and hardware overhead. Therefore, it can evade the modern Trojan detection technique such as power-analysis. The Trojan is only activated by the implanter by a sequence three plaintexts which is chosen in such a way that it reduced the possibility of accidental activation of the Trojan. This work also brings out a new challenge of trusted design flow where there are several untrusted steps.

Finally, the work focuses on developing suitable counter-measures against fault attacks. There are some counter-measures which require huge area overhead. On the other hand the counter-measures proposed in literature are suitable for single byte faults. However with multi byte faults, the counter-measures have large overhead, this motivates us to study low-cost and low-overhead counter-measures against fault attacks on AES.

Soma Saha
Email: somasaha45@yahoo.co.in
Joined the department in: July 2009

***Soma Saha** received a B.Tech. degree in Computer Science & Engineering from University College of Science & Technology, University of Calcutta, Kolkata in 2007, and an M.Tech. degree in Computer Science & Engineering from University College of Science & Technology, University of Calcutta, Kolkata in 2009. From July 2007 till 14th July 2009, she attached with Maharaja Manindra Chandra College, University of Calcutta, Kolkata, as a Guest-Lecturer in Deptartment of Computer Science. Since 22nd July 2009, she has been a research scholar in the department of Computer Science & Engineering in IIT Kharagpur. Her research interests are in the areas of Multi-Objective Combinatorial Optimization and Evolutionary programming.*

Supervisor: Prof. Rajeev Kumar

Bounded-Diameter MST solutions with Hybridization of Multi-Objective EA

Minimum Spanning Tree (MST) has many applications, particularly in electrical, communication and transport networks, as well as in distributed mutual exclusion algorithms and in data compression. Much research has been performed on MST problem with a particular diameter constraint, which is known as Bounded Diameter Minimum Spanning Tree (BDMST)/ Diameter Constraint Minimum Spanning Tree (DCMST) problem. Deterministic polynomial time algorithm exists for BDMST problem having diameter bound at most 3. But, it is an NP-hard problem within diameter range 4 <= diameter,D < |V|-1; where V is the set of vertices of a complete graph G. Existing well-known deterministic heuristics for BDMST problem are: One Time Tree Construction (OTTC) and Iterative Refinement (IR), Center Based Tree Construction (CBTC), Randomized Greedy Heuristics (RGH) or Randomized Tree Construction (RTC), and Center Based Recursive Clustering (CBRC). There are well-known stochastic methods that include evolutionary algorithms (EA), like edge-set encoding EA (JR-ESEA), permutation-coded EA (PEA) and among other well-known approaches are Variable Neighbourhood Search (VNS), Modified Hybrid Genetic Algorithm (MHGA). Singh and Gupta [1] suggested two improvements over CBTC and RGH and PEA, named as CBTC-I, RGH-I and PEA-I, to get better solutions for BDMST problem.

We recast single-objective MST problem with a constraint on diameter to a bi-objective MST problem with two conflicting objectives (minimizing weights and minimizing diameters) and generate Pareto front for entire range of diameter 2 to |V|-1. Some work has been done using evolutionary programming (EP) by Kumar et al. [2] and quality performance of different heuristics and EA for bi-objective MST problem has been done by Kumar and Singh [3]. Previous works on bi-objective MST problem has been motivated on improving Pareto front from randomized initial population for Euclidean instances only. In our work, we have analyzed the nature of solutions generated by recasting BDMST heuristics to solve bi-objective MST problem on entire diameter range for non-Euclidean instances. We have provided an hybridized Multi-Objective Evolutionary Algorithm (MOEA) to improve the Pareto front solution-set obtained from each heuristics for both Euclidean and non-Euclidean instances.

Hybridized MOEA approach:

- **Step 1:** Initialization population with the Pareto front solutions which are generated from particular heuristic.
- **Step 2:** Randomly selection of parents to perform crossover/recombination operation with crossover probability p_c
- **Step 3:** Randomly selection of individuals from population with mutation probability p_m to perform mutation operation.
- **Step 4:** Perform selection operation on union of initial population, individuals from crossover operation and individuals from mutation operation and creation of population for next generation.
- **Step 5:** Repeat Step 2 to Step 4 until termination condition reaches.

We have adapted well-known crossover [4] and mutation operators [4] to reduce the complexity of basic genetic operators.

We have considered a reference Pareto front from all improved Pareto fronts obtained for each heuristics for a problem instance. Each Pareto front obtained using MOEA for Euclidean instances tends nearer to reference Pareto front than Pareto fronts obtained from heuristics. RGH provides better result for Euclidean instance within a very small range of diameter; on applying EA, the solutions obtained cover a wide range of diameter for Euclidean instances. Similar improvement is observed for non-Euclidean instances; but, here, OTTC and CBTC generate superior solutions with a very small drift across the complete range of diameter. On applying EA, for each heuristics, the solution set is improved. Thus, we conclude that hybridization of heuristic solutions with EA improves the solutions across the complete range of diameter which in turn improves the BDMST solutions for a particular diameter constraint.

References

1. Singh and A. K. Gupta. Impoved Heuristics for the Bounded Diameter Minimum Spanning Tree Problem. Journal Soft Computing, 11: 911-921, 2007.
2. R. Kumar and B. K. Bal and P. I. Rockett. Multiobjective Genetic Programming Approach to Evolving Heuristics for the Bounded Diameter Minimum Spanning Tree Problem. GECCO'09, ACM.
3. Rajeev Kumar and Pramod Kumar Singh. On quality performance of heuristic and evolutionary algorithms for biobjective minimum spanning trees. GECCO '07, 2007, 2259, ACM Press.

Soumen Bag
Email: soumen@cse.iitkgp.ernet.in
Joined the department in: July 2008

***Soumen Bag** received B.E. degree in Computer Science & Engineering from NIT, 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 & Engineering. He received M.Tech. degree in Computer Science & Engineering (specialization in Information Technology) from NIT, Durgapur in 2008. Since July 2008, he has been a research scholar in the department of Computer Science & Engineering in IIT Kharagpur. His research interests are in the areas of OCR for Indian Scripts, Document Image Analysis, Image Processing, and Pattern Recognition.*

Supervisors: Prof. Partha Bhowmick and Prof. Gaurav Harit (IIT Rajasthan)

Recognition of Handwritten Bengali Characters Using Skeletal Convexity and Dynamic Programming

The main challenge in recognizing handwritten characters is to handle large-scale shape variations in the handwriting of different individuals. We see from the past several decades that different types of features are proposed for printed and handwritten OCR systems for Indian languages. But the performance of handwritten OCR still needs a lot of improvement. From this point of view, we propose a novel method based skeletal convexity and dynamic programming for Bengali handwritten character recognition. Structural shape of a character is described by different skeletal convexities of character strokes irrespective of viewing direction on 2D plane. Such skeletal convexity acts as an invariant feature for character recognition.

Given a scanned document page, we first binarize it using Otsu’s algorithm. Currently we are working with isolated character images. Before extracting the structural shape, character images are converted to thinned (i.e., single pixel thick) curve segments. But to retain the proper shape of thinned character images is a big challenge. Here we consider a medial-axis based thinning strategy for performing character skeletonization as a preprocessing. For noisy images, the proposed medial-axis based thinning results in undesired small concave and convex regions. To solve this problem, we apply a straight line approximation method on thinned images. The approximation results often contain deviation of thinned images at the junction points. So, to preserve the true shape at the junction points during approximation, we perform junction point refinement.

After applying straight line approximation method on thinned images, we get a set of approximation points $V = \{p_1, p_2, ..., p_n\}$ and a set of edges $E = \{e_1, e_2, ..., e_n\}$ connecting approximation points according to the structural shape of character images. Now, we traverse the graph, starting from an end point, as a cyclic manner to get a sequence of visited points. Next, we detect the concavity and convexity of all these points (except the start and end points of traversal) by calculating the value of twice the signed area of a triangle (Eq. 1) formed by the point p_i (x_i, y_i) and its two adjacent points, $p_{i-1}(x_{i-1}, y_{i-1})$ and $p_{i+1}(x_{i+1}, y_{i+1})$.

$$\Delta(p_{i-1}, p_i, p_{i+1}) = \begin{vmatrix} 1 & 1 & 1 \\ x_{i-1} & x_i & x_{i+1} \\ y_{i-1} & y_i & y_{i+1} \end{vmatrix} \tag{1}$$

If $\Delta(\cdot)$ yields a negative value, then the point p_i has the concave property and is marked as *L*. If the value is positive, then p_i has a convex property and is marked as *R* (Fig. 1). If the value is equal to 0, then the point p_i has the same property of its previous point p_{i-1}. Finally, longest common subsequence (LCS) matching is used on *LR* sequence for character recognition. This dynamic programming approach has been found to be useful in some contemporary online handwriting recognition systems which has motivated us to apply it in our off-line system.

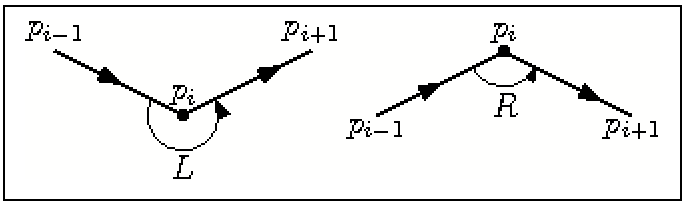


Fig. 1: *Detection of concavity and convexity of a point with respect to its neighbor points. left: Concave shape; right: Convex shape.*

For experimental purpose, we have taken the handwritten Bengali character database of ISI, Kolkata. The database contains 50 Bengali basic characters with wide variation. We have taken 10 different shapes for each character from the database. At first, we have prepared a set of prototypes which contains the *LR* sequences of all printed Bengali basic character images. This prototype set acts as a ground truth for handwritten character recognition. Now, the handwritten characters are recognized using the LCS matching with the prototypes. Preliminary results demonstrate the efficacy of our approach. But this method is not performing well for few particular characters. In future, we shall extend our work to improve the accuracy of recognition and to make it applicable to character classification for handwritten Bengali OCR system.

Soumyadip Bandyopadhyay
Email: soumyadip@cse.iitkgp.ernet.in
Joined the department in: January 2009

***Soumyadip Bandyopadhyay** received a B.Tech. degree in Computer Science and Engineering from Bengal Institute of Technology, Kolkata in 2008. Since January 2009, he has been a research scholar in the department of Computer Science & Engineering in IIT Kharagpur. His research interests are in the areas of Formal Verification of Embedded Systems.*

Supervisors: Prof. Chitta Ranjan Mandal and Prof. Dipankar Sarkar

Use of PRES$^+$ Models in High-Level Verification

We focus on some aspects related to modeling and formal verification of embedded systems. Many models have been proposed to represent embedded systems [1] [2]. These models encompass a broad range of styles, characteristics, and application domains and include the extensions of finite state machines, data flow graphs, communication processes and Petri nets. In this report, we have used PRES$^+$ models (Petri net based Representation for Embedded Systems) as extension of classical Petri net models that capture concurrency and timing behaviour of embedded systems; it allows systems to be representative at different levels of abstraction and improves expressiveness by allowing the token to carry information [3]. This modeling formalism has a well defined semantics so that it supports a precise representation of systems. As a first step, we have taken an untimed PRES$^+$ model which captures all the features of PRES$^+$ model except the time behaviour.

A typical synthesis flow of complex systems like VLSI circuits or embedded systems comprises several phases. Each phase transforms/refines the input behavioural specification (of the systems to be designed) with a view to optimize time and physical resources. Behavioural verification involves demonstrating the equivalence between the input behaviour and the final design which is the output of the last phase. In computational terms, it is required to show that all the computations represented by the input behavioural description, and exactly those, are captured by the output description.

Modeling using PRES$^+$, as discussed above, may be convenient for specifying the input behaviour because it supports concurrency. However, to the best of our knowledge there is no equivalence checking method reported in the literature for PRES$^+$ models. As a first step, therefore, we seek to devise an equivalence checking procedure for PRES$^+$ models by translating them to FSMD models and then use FSMD equivalence checker [4]. We demonstrate the working of the algorithm on a real life example. For this purpose, we have used non-pipelined and pipelined version of an aluminum extraction plant controller.

As a part of the future work, we intend to evolve a direct method for equivalence checking between of two PRES$^+$ models. We also intend to generalize FSMD models to timed FSMD models and then conversion from PRES$^+$ models to timed FSMD models. We will also devise equivalence checking algorithm between two timed FSMD models.

References

1. S. Edwards, L. Lavagno, E. A. Lee, and A. Sangiovanni-Vincentelli, “Design of embedded systems: Formal models, validation, and synthesis,” in Proceedings of the IEEE, pp. 366–390, 1997.
2. P. Eles, K. Kuchcinski, Z. Peng, A. Doboli, and P. Pop, “Scheduling of conditional process graphs for the synthesis of embedded systems,” in DATE ’98: Proceedings of the conference on Design, automation and test in Europe, (Washington, DC, USA), pp. 132–139, IEEE Computer Society, 1998.
3. L. A. Cort´es, P. Eles, and Z. Peng, “Verification of embedded systems using a petri net based representation,” in ISSS ’00: Proceedings of the 13th international symposium on System synthesis, (Washington, DC, USA), pp. 149–155, IEEE Computer Society, 2000.
4. C. Karfa, D. Sarkar, C. Mandal, and P. Kumar, “An equivalence-checking method for scheduling verification in high-level synthesis,” IEEE Trans. on CAD of Integrated Circuits and Systems, vol. 27, no. 3, pp. 556–569, 2008.

Soumyajit Dey
Email: soumyad@cse.iitkgp.ernet.in
Joined the department in: July 2007

***Soumyajit Dey** received B.E. degree in Electronics and Telecommunication Engineering from Jadavpur University, Kolkata in 2004, and M.S. degree in Computer Science 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 IIT Kharagpur. His research interests are in the areas of Formal Methods for Modeling and Verification of Embedded Systems.*

Supervisor: 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 models (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 system specifications which is possible using formalisms like Kahn Process Network (KPN), Timed Automata (TA), Synchronous Dataflow Graphs (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.

Further, we provide a TSM based model of the European Train Control System (ETCS) protocol and perform the safety property verification by applying KAT rules over the Kleene expressions of the actor network.

References

1. Benveniste, B. Caillaud, L. Carloni, P. Caspi and A. Sangiovanni-Vincentelli, “Composing Heterogeneous Reactive systems”, ACM Trans. Embedded Computing Systems, Vol. 7, No. 4, pp 1-36, 2008.
2. Dexter Kozen, “Kleene algebra with tests”, ACM Trans. Programming Languages and Systems, Vol. 19, No. 3, pp. 427-443, 1997.

Sourav Kumar Dandapat
Email: sdandapat@cse.iitkgp.ernet.in
Joined the department in: July 2009

***Sourav Kumar Dandapat** received a B.E. degree in Computer Science from Jadavpur University in2002, and an M.Tech. degree in Computer Science from IIT Kharagpur in 2005. From July 2005 till November2007, 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, hehas 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

Applications and solutions for next generation Wireless Internet

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 Kleenesemirings. Such a characterization helps in equational reasoning on heterogeneous specifications using the axioms of Kleene algebra.

Association control scheme for wireless mobile environment

Wireless clients 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 AP with the strongest signal around it. The main drawback in RSSI based technology is that the global parameters are not considered during association, hence effective strategy to handle skewed geographical distribution of devices (thereby ensuring fairness) cannot be devised. However, in today’s enterprise WLANs, multiple APs are getting connected to a central controller through high speed wired backbone. As a result, modern networks are becoming semi-centralized through hybrid wired-wireless architecture that offers new opportunities to redesign protocols for future wireless. Hence, there is a need to develop smart association control schemes which will ensure higher admission along with fairness, exploiting the global view of the APs. This is particularly pronounced in light of enterprise WLANs shifting to the single wide-channel mode (proposed by Meru Networks) to reduce the problems of interference management and frequent handoff. Association control is likely to play a key role in such environments. So, the broad objectives of our research can be summarized as follows – (a) Develop an AP-guided association control strategy that exploits the global view of APs for association decision, and (b) Maximize the number of connections admitted while maintaining fairness in bandwidth allocation.

Collaborative Download Exploiting Social Behaviour

Although association control is important to maximize the number of devices admitted, wireless technology still suffers from low data rates and high costs of accessing the Internet. The Wi-Fi technology has not proliferated enough to compensate for this cost. Hence an important line of research is in exploring a protocol possibility whereby wireless devices can communicate among themselves and download collaboratively from the internet. In a simplified form, each participating device may download a part of the targeted object from the Internet which is subsequently exchanged. It is seen that that movements of people follow certain social patterns-people meet/interact with peers who have similar type of interest, and they download similar type of (especially temporally relevant eg. stock prices or latest hit songs) files. There are however, a number of challenges involved in this work, such as group formation (ad-hoc network formation with user's consensus), scheduling and work distribution (based on data rate, computation power and incurring cost), implementation of proper incentive schemes (challenges imposed due to highly dynamic behavior of mobile nodes), data exchange among users using Bluetooth, Wi-Fi technology and so on. We plan to develop a protocol that will enable such collaborative/aggregated downloads among social groups of human individuals. We hope to make important contribution along with increasing number of researchers who are working in this direction to overcome the hurdles involved.

Srobona Mitra
Email: srobona@cse.iitkgp.ernet.in
Joined the department in: March 2007

***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 and Post-Silicon Verification.*

Supervisors: Prof. Pallab Dasgupta and Prof. Partha Pratim 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 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. 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, and a design change due to a bug fix, both upwards and downwards along the design hierarchy, so that the entire framework gets updated with the implications of the new property and bug fix.

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

Subhadip Kundu
Email: subhadip@iitkgp.ac.in
Joined the department in: July 2010

***Subhadip Kundu** received Bachelor of Technology Degree (B.Tech) from West Bengal University of Technology in Electronics and Communication Engineering in the year 2007. He received MS degree from Indian Institute of Technology Kharagpur in 2010 from Electronics and Electrical Communication Department. His MS research topic was Low Power Testing. Currently, he is pursuing PhD from Computer Science and Engineering Department, Indian Institute of Technology Kharagpur. His current areas of research are: Fault diagnosis in digital VLSI system and Power aware testing. He has published more than 10 international conference papers and journals in this domain.*

Supervisor: Prof. Indranil Sengupta and Prof. Santanu Chattopadhyay (ECE Dept.)

Fault Diagnosis in Digital Systems

Diagnosis is the methodology to identify the reason behind the failure of manufactured chips. This is particularly important from the yield enhancement viewpoint. Though many of the existing diagnosis algorithms are based upon the *cause-effect analysis* using a *fault dictionary*, the approach is restrictive in the sense that it relies on a particular fault model, and thus not suitable for different types of faults that may occur in today's IC manufacturing process. In this light, the *effect-cause analysis* based approaches appear to be more viable. In such an approach, for the *failing vectors* (test vectors for which at least one primary output differs from the fault-free case), probable faulty signals are identified. Based on a relative ranking between these signals, the final fault locations are predicted. However, most of the existing approaches are centered on identifying upto four simultaneous faults only. Another major problem in fault diagnosis occurs in scan-chain diagnosis. These failures often account for upto 50% of chip failures. Therefore, scan chain failure diagnosis is important to effective scan-based testing.

The followings are the objective of the work:

- Using effect-cause analysis to locate multiple faults in combinational logic
- Finding suitable metric to compare the diagnostic capability between test sets
- Incorporating diagnostic test pattern generation (DTPG) to enhance diagnostic quality of a test pattern set

- Modifying circuit structure to have better diagnosability

Almost all the conventional fault diagnosis method simulate one fault (among the candidate faults) at a time and based on the number of failed patterns the fault can explain, a ranking is proposed for all candidate faults. But, a single fault injection cannot manifest the effect of multiple faults that are present in the actual failed circuit. Thus, in this work, we have proposed a fault diagnosis algorithm based on multiple fault simulation. Since, the number of faulty sites is unknown; multiple fault simulation algorithms are inherently exponential in time. So, to cover this exponential search space, we have used a Particle Swarm Optimization (PSO) algorithm for finding suitable solutions. Initially, a list of possible fault candidates have been found out by critical path tracing from each failing POs and taking a union of them. Now, the initial particles of PSO are chosen at random from the possible faulty sites and the number of faults in each particle is also varying. So, each particle is a set of faults with varying cardinality. These particles are then evaluated based on how many patterns (both pass and fail) they can explain. The final result is given by a set of tuple of faults which have successfully explained the entire pattern set. Theoretically, our algorithm can diagnose successfully up to any number of faults present in the circuit.

The flow of our fault diagnostic algorithm is as follows:

- First we start with a failing PO (for a particular failed pattern) and backtrace its fan-in cone using a critical path tracing method. The same process is repeated for all the failing POs for all failing patterns.
- Since we are considering multiple faults, union of all the fan-in cones is taken. All the faults obtained by this method are considered as the candidate faults and given as an input to the PSO.
- PSO also takes the failed pattern set with the faulty responses (from the tester) as well as the pass patterns as its other inputs. The algorithm generates multiple sets of faults having same fitness as the best particle.
- The results produced by PSO needs to go for further pruning.
- The final result produced by our method is a ranked list of possible faults. We also give the best solution provided by the PSO.

Subhankar Mukherjee
Email: soumyad@cse.iitkgp.ernet.in
Joined the department in: July 2007

***Subhankar Mukherjee** received a 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 & Engineering in Indian Institute of Technology Kharagpur. His research interests are in the area design verification.*

Supervisors: Prof. Pallab Dasgupta and Prof. Siddhartha Mukhopadhyay

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. A mixed-signal assertion 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 SystemVerilog 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.*

Subhasish Dhal
Email: subhasis.rahul@gmail.com
Joined the department in: December 2009

***Subhasish Dhal** received a B. Sc (H) degree in Computer Science from Vidyasagar University, Midnapore in 2002, and a MCA degree from NIT Durgapur in 2005. He also has received an M. tech degree in Computer Sc. and Engineering from NIT Rourkela in 2009. From August 2005 till August 2007 he worked in Asutosh College, Kolkata as a lecturer (contractual) and from August 2009 till December 2009 he worked in IE & IT, Durgapur as a lecturer. Since December 2009, he has been a research scholar in the department of Computer Science & Engineering in IIT Kharagpur. His research interests are in the areas of Security in RFID and Mobile Networks.*

Supervisor: Prof. Indranil Sengupta

Authentication in RFID Communication

RFID technology defeats bar code reader in respect to the efficient reading capability and it does not need line of sight constraint. Therefore, the objects are tagged by RFID device which contains the information related to the object. Sometimes, the information related to an object is such that the RFID device is unable to contain the whole. For those objects, a backend data base is used to keep the relevant information. The tag only keeps the key information. RFID reader reads the key information and consults with the backend database for the detail information. The information related to the objects may be secret and hence the access needs to be secure. Therefore, security such as privacy, authentication, integrity etc is the basic requirements for these devices. We are focusing on the authentication issues. Use of multiple RFID tags in an object increases the detection rate in comparison to the single tagged object. Since, more than one tags are involved in the same object, the security such as authentication scheme needs to be revised. Our objective is to build a lightweight authentication scheme which will give maximum security in respect to the authentication to the RFID communication.

Sudip Roy
Email: sudipr@cse.iitkgp.ernet.in
Joined the department in: October 2009

Sudip Roy *received B.Sc. (Honors) in Physics and B.Tech. in Computer Science and Engineering from the University of Calcutta, Kolkata, India in 2001 and 2004, respectively, and M.S. in Computer Science and Engineering from Indian Institute of Technology Kharagpur, India, in 2009. He is currently pursuing his Ph.D. in Computer Science and Engineering at Indian Institute of Technology Kharagpur, India. His areas of research interest include algorithms for computer-aided design and testing of digital VLSI circuits and digital microfluidic biochips.*

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

Automated Dilution of Biochemical Fluids in Digital Microfluidic Biochips

Microfluidic-based biochips are soon revolutionalizing clinical diagnostics and other biochemical laboratory procedures to meet the challenges of healthcare cost for cardiovascular diseases, cancer, diabetes, and global HIV crisis, etc. [1-3]. 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. [1-4]. 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 [1-3].

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 and promising 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) [1-3]. 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 [1-4].

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]. Several combinatorial and geometric optimization problems arise during the computer-aided design and testing of such chips [4]. 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.

A key challenge in design automation of DMF biochips is to carry out the dilution and mixing of fluids onchip. An existing algorithm for dilution/mixing of fluids minimizes the number of mixing steps to achieve the target concentrations (Thies et al., Natural Computing, 2008). However, in a bioprotocol, waste droplet handling is cumbersome and the number of waste reservoirs onchip should be minimized to use limited amount of sample and expensive reagents, and hence to reduce the cost of the biochip. Thus, waste reduction is crucial during dilution/mixing of fluids. First, we present an optimization algorithm that significantly reduces waste droplet production compared to the prior work. Next, we design another improved algorithm that optimizes the usage of intermediate droplets generated during dilution process to reduce input demands and production of waste droplets. An integrated scheme is presented for choosing the best waste-aware dilution algorithm among these three methods for a target concentration. To realize a dilution-on-chip we design an architectural layout of electrodes for DMF biochips consisting of two rotary mixers and some storage electrodes.

References

1. K. Chakrabarty, and F. Su, “Digital Microfluidic Biochips: Synthesis, Testing and Reconfiguration Techniques”. CRC Press, 2007.
2. R. B. Fair, “Digital Microfluidics: Is a True Lab-on-a-Chip Possible?” Microfluidics and Nanofluidics, Vol. 3, March 2007.
3. M. Abdelgawad, and A. R. Wheeler, “The Digital Revolution: A New Paradigm for Microfluidics”, Advanced Materials, Vol. 21, 2009.
4. K. Chakrabarty, and J. Zeng, “Design Automation for Microfluidic-Based Biochips”, ACM JETC, Vol. 1, No. 3, October 2005.

Sudipta Saha
Email: sudipta.saha.22@gmail.com
Joined the department in: July 2010

***Sudipta Saha** received B. E. degree in Computer Science & Technology, from Bengal Engineering College (at present known as Bengal Engineering and Science University), Shibpore in 2002 and MTech degree from Indian Institute of Technology (IIT), Kharagpur in 2008. He worked as 'Senior Lecturer' in a college affiliated under West Bengal University of Technology (WBUT). He also served as 'Associate Member of Technical Staff' in Magma Design Automation Pvt. Ltd, Noida. In 2008, he joined PowerSys Technologies Pvt. Ltd., a start up organization established by two senior faculty members of IIT Kharagpur. Since July 2010, he has been a research scholar in the department of Computer Science & Engineering, IIT Kharagpur. His research interests are in the areas of Computer Network and Bioinformatics.*

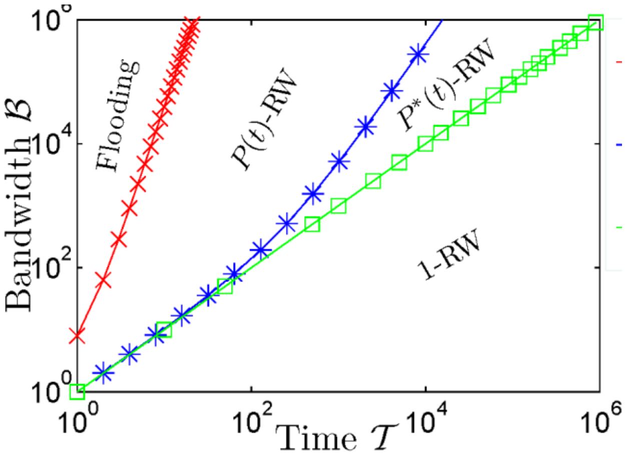
Supervisor: Prof. Niloy Ganguly

Coverage Maximization under Resource Constraint in Unstructured Networks
Significance of Coverage algorithms

Search and dissemination of information – are one of the most frequently used tools in the Internet. And the same is true for any general purpose distributed system. The algorithms which are responsible for performing these operations are termed as 'coverage algorithm'. There has always been a great requirement for covering as much area of the network as possible in lesser time. For structured networks, because of their implicit structure, there are good solutions to this problem of covering. However, due to lack of any sense of global idea or any information related to the structure of the network, the covering algorithms available for unstructured networks are not efficient. But, unfortunately, because of its inherent resilience against most frequently performed operations like - joining of new node in the network, leaving of an existing node from the network etc, most of the real world networks are unstructured or semi-structured in nature. The problem is more severe, when the available resource which is consumed for transmission of message packets – is limited. For example, in the wireless sensor network, the battery power is very limited. Available network bandwidth (B) – can also be a big constraint on these algorithms. In addition, time (T) can always be thought of as a constraint. Consequently, the problem of coverage maximization under resource constraints – becomes a significant issue in these areas.

Concept of Optimality in terms of Coverage

For unstructured networks, the easiest and most inefficient naïve approach is Flooding. It wastes most of the available resources except time. So, if we define the coverage as a function of available bandwidth B and Time T, then flooding is optimal in a specific region of the arena. In unstructured networks, if we do not want that much speed of search, then we have another option – which is based on symmetric random walk. When a single random walker (1-RW) is used for the same purpose, it wastes the least amount of bandwidth, but is excessively slow. So, this is also optimal in the sense that it wastes least bandwidth thereby giving highest coverage. But the concept of optimality is not at all clear in the region between flooding and the single random walker. There are varieties of algorithms available which are either based on random walk or on flooding. A variation of proliferated random walk based scheme (P*(t)-RW) was derived by Subrata Nandi [1], a research scholar in the department of Computer Science & Engineering, IIT Kharagpur, in 2010, which addresses the definition of optimality in the region between flooding and 1-RW. It gives the idea that, it is possible to achieve the same coverage as that of a single random walker but with much faster rate. Therefore, it explored optimality in some more region of the spectrum. But still the region beyond the P*(t)-RW lacks the concept of optimality. It is unknown that whether one can achieve the same coverage as that of a 1-RW while consuming the same bandwidth and taking lesser time than P*(t)-RW. The following figure shows this basic problem. P(t)-RW, a derived variation of P*(t)-RW, falls in the specified region in the figure. But it is not clear that this variation or any other existing algorithm is optimal or not in that region.



Our Approach:

We started redefining the concept of optimality in terms of many other factors. We found that there is a lack of metrics that can judge real goodness of a given coverage algorithm. This actually varies depending on the real requirement. So, we first devised few target functions and measured some important metrics which we used for the purpose of testing the existing algorithms for coverage. We observed that careful modification of proliferation techniques can give much better performance. We also observed that not only the time, but also the distance is a significant concept in designing different proliferation techniques. At present we are doing experiments on different time and distance based proliferation techniques. Our future works include defining real profile of the impact of distance in the proliferation as well as turning the theoretical work into a practical work and thereby improving the algorithms in real life networks.

Reference

1. S. Nandi, L.Brusch, A. Deutsch, and N. Ganguly. Coverage-maximization in networks under resource constraints. Phys. Rev. E, 81(6):061124, Jun 2010.

Sumanta Pyne
Email: sumantapyne@gmail.com
Joined the department in: December 2009

Sumanta Pyne *received a B.Tech. degree in Computer Science from Meghnad Saha Institute of Technology, Kolkata in 2005, and an M.E. degree in Computer Science from Bengal Engineering and Science University, Shibpur in 2009. From July 2005 till January 2006, he worked in Hi-Q Solutions, Kolkata, as a programmer. From January 2006 till June 2007 he was a lecturer at Techno India College of Technology, Kolkata. Since December 2009, he has been a research scholar in the department of Computer Science & Engineering in IIT Kharagpur. His research interests are in the areas of Power Aware Software.*

Supervisor: Prof. Ajit Pal

Optimization of Power Aware Software Prefetching

Software prefetching is a performance-oriented optimization technique, which is generally used to reduce the gap between processor speed and memory access speed. When software prefetching is applied to memory-intensive benchmark programs, the performance improves with higher power consumption. The present work provides a mechanism to transform a program with software prefetching to its power aware equivalent. This is done by executing the software prefetching program at different voltage-frequency pairs. Besides reducing the power, the performance has been improved by adjusting the prefetch distance. Xeemu-Panalyzer simulator is used to evaluate the present work. Experimental results of the proposed scheme guarantees that performance improvement of software prefetching program is possible at the cost of less power consumption. The proposed work can enable a compiler to generate power aware software prefetching program.

Power/Energy saving is vital issue not only for embedded and portable devices, but also for servers, personal, mainframe and super computers. This has lead to design compilers for power aware code generation [1]. It has been observed some code optimization techniques lead to performance gain as well as saves energy. But it is not true for all. Software prefetching is such a technique. Software prefetching [6, 2] is a technique of prefetching data from main memory to cache in advance, well before the processor needs the data for computation. Software prefetching is done by means of prefetch instructions supported by the instruction set of the processor. Software prefetching eliminates cache miss causing improvement in performance. It also increases power consumption. Dynamic Voltage Frequency Scaling (DVFS) [5] is an effective technique for low power but it also degrades the performance of the system. In [3] Agarwal et al have proposed the idea of low power software prefetching using DVFS, without degrading the performance. While Chen et al in [4] shows DVFS with adjustment of prefetch distance can provide power reduction as well as performance improvement. The present work formulates the problem of power reduction with performance gain as two optimization problems, a single objective optimization problem and a multi-objective optimization problem. The solution of these optimization problems will guide the software prefetching program to achieve higher performance at the cost of minimum power consumption.

References

1. Vivek Tewari, Sharad Malik and Andrew Wolfe“. A Document Preparation System”. In the Proceedings of the 1994 Symposium on Low-Power Electronics, San Diego, CA, October 1994.
2. Todd C. Mowry.Tolerating. “Latency through Software-Controlled Data Prefetching”. Doctor dissertation, Standford University, March 1994.
3. Deepak N. Agarwal, Sumitkumar N. Pamnani, Gang Qu, and Donald Yeung. “Transferring Performance Gain from Software Prefetching to Energy Reduction”. In Proceedings of the 2004 International Symposium on Circuits and Systems (ISCAS2004). Vancouver, Canada. May 2004.
4. Juan Chen, Yong Dong, Huizhan Yi, and Xuejun Yang. “Power-Aware Software Prefetching”. ICESS 2007, LNCS 4523, pp.207218. Springer-Verlag Berlin Heidelberg 2007.
5. Fen Xie, Margaret Martonosi and Sharad Malik. “Intraprogram Dynamic Voltage Scaling: Bounding Opportunities with Analytic Modeling”. ACM Transactions on Architecture and Code Optimization. Vol.1, No.3, September 2004. pages 323-367.
6. D. Callahan, K. Kennedy, and A. Portereld. “Software Prefetching”. In the Proceedings of the 4th International Conference on Architectural Support for Programming Languages and Operating Systems, Santa Clara, CA, April 1991.

Tirthankar Dasgupta
Email: iamtirthankar@gmail.com
Joined the department in: January 2010

Tirthankar Dasgupta *received a B.E. degree in Information Technology from MCKV Institute of Technology, Kolkata in 2003, and an MS degree in Computer Science from Indian Institute of Technology, Kharagpur in 2009. From January 2009 till December 2009, he worked in Society for Natural Language Technology Research, Kolkata as a Researcher. Since January 2010, he has been a research scholar in the department of Computer Science & Engineering in IIT Kharagpur. His research interests are in the areas of Natural Language Processing, Cognitive Science, psycholinguistics and Assistive Technology.*

Supervisor: Prof. Anupam Basu

Towards a Computational Model for the Organization and Access of Mental Lexicon: A Journey with Bangla Words

Understanding the organization of the *mental lexicon* is one of the important goals of cognitive science. *Mental lexicon* refers to the representation of the words in the human mind and the various associations between them that help fast retrieval and comprehension of the words in a given context. Words are known to be associated with each other at various levels of linguistic structures namely, orthography, phonology, morphology and semantics. However, the precise nature of these relations and their interactions are unknown and very much a subject of research in psycholinguistics. A clear understanding of these phenomena will not only further our knowledge of how the human brain processes language, but also help in developing apt pedagogical strategies and find applications in natural language processing.

One of the key questions that psycholinguists have been investigating for a long time and debating a lot about is the mental representation and access mechanisms of polymorphemic words: whether they are represented as a whole in the brain or are understood by decomposing them into their constituent morphemes. That is to say, whether a word such as "*unimaginable*" is stored in the mental lexicon as a whole word or do we break it up "*un-*" , "*imagine*" and "*-able*", understand the meaning of each of these constituent and then recombine the units to comprehend the whole word. Such questions are typically answered by designing appropriate priming experiments or other lexical decision tasks. The reaction time of the subjects for recognizing various lexical items under appropriate conditioning reveals important facts about their organization in the brain.

There is a rich literature on organization and lexical access of polymorphemic words where experiments have been conducted mainly for English, but also Hebrew, Italian, French, Dutch, and few other languages (Frost et al., 1997; Marslen-Wilson et al. 1994). However, we do not know of any such investigations for Indian languages, which are morphologically richer than many of their Indo-European cousins. On the other hand, several cross-linguistic experiments indicate that mental representation and processing of polymorphemic words are not quite language independent (Taft, 2004). Therefore, the findings from experiments in one language cannot be generalized to all languages making it important to conduct similar experimentations in other languages. Bangla, in particular, features stacking of inflectional suffixes (e.g., *chhele* + *TA* + *ke* + *i* "to this boy only"), a rich derivational morphology inherited from Sanskrit and some borrowed from Persian and English, an abundance of compounding, and mild agglutination.

The primary objective of this research is to understand the organization of the Bangla mental lexicon at the level of *morphology*. Our aim is to determine whether the mental lexicon decomposes morphologically complex words into its constituent morphemes or does it represent the unanalyzed surface form of a word. We apply the cross modal repetition priming technique to answer this question specifically for derivationally suffixed polymorphemic words of Bangla. We observe that morphological relatedness between lexical items triggers a significant priming effect, even when the forms are phonologically unrelated. On the other hand, phonologically related but morphologically unrelated word pairs hardly exhibit any priming effect. These observations are similar to those reported for English and indicate that derivationally suffixed words in Bangla are accessed through decomposition of the word into its constituent morphemes.

Further analysis of the reaction time and error rates per word and per subject reveal several interesting facts such as (a) apart from usage frequency, word length and presence of certain orthographical features also affect the recognition time of a word, and (b) certain derivational suffixes inherited from Sanskrit, which usually make the derived word phonologically or semantically opaque, do not trigger priming; this indicates that these morphological relations are no longer recognized or internalized by the modern Bangla speakers. These and similar other observations make us believe that understanding the precise nature of the mental representation of morphological processes in Bangla (as well as other Indian languages) is a challenging and potent research area that is very little explored.

Reference

1. Frost, R., Forster, K.I., & Deutsch, A. (1997). What can we learn from the morphology of Hebrew? A masked-priming investigation of morphological representation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23,* 829–856.
2. Marslen-Wilson, W.D., Tyler, L.K., Waksler, R., & Older, L. (1994). Morphology and meaning in the English mental lexicon. *Psychological Review*, *101*, pp. 3–33.

MS Scholars

List of Current M.S. Scholars

Animesh Srivastava
Anup Kumar Bhattacharya
Binanda Sengupta
Biswajit Das
Biswanath Barik
Biswanath Saha
Chandan Misra
Debjit Pal
Debmalya Sinha
Partha De
Praloy Kumar Biswas
Pramit Roy
Prosenjit Dhole
Rajdeep Mukherjee
Ramji Nagariya
Ritwika Ghose
Sandipan Mandal
Satrajit Ghosh
Sirsendu Mohanta
Sourasis Das
Sourya Bhattacharyya
Souvik Bhattacherjee
Sujoy Sinha Roy
Sumit Das
Suprabhat Das
Swarnendu Biswas

Research Abstracts
(MS)

Animesh Srivastava
Email: asrivastava@cse.iitkgp.ernet.in
Joined the department in: September 2009

***Animesh Srivastava** received a B.Tech.degree in Computer Science & Engineering from Haldia Institute of Technology, Haldia in 2007. From July 2007 till September 2009, he worked in Wipro Technologies, Bangalore, as a Project Engineer. Since September 2009, he has been a research scholar in the department of Computer Science & Engineering in IIT Kharagpur. His research interests are in the areas of Resilience of real-world P2P Networks.*

Supervisor: Prof. Niloy Ganguly

Stability Analysis Real-World P2P Networks

Popular peer-to-peer networks like Gnutella, Kazaa are increasingly subjected to various kinds of attacks like Denial of Service attack (DoS), DDoS attack, Eclipse attack, Sybil attack etc. All these attacks try to interrupt the network-wide peer communication by disrupting the activities of the highly connected (resourceful) nodes. Besides, the continuous churn of the constituent nodes may also lead to interruption in the network-wide communication. Analytical work predicting the outcome of such churn and attack on large dynamic networks has been studied in depth, in the last decade. The results are primarily based upon the concept of percolation theory whereby the relation between component size and attack is established. These works have been successfully extended in the domain of p2p networks. However, it has been observed that these theories work perfectly for random networks and fail when applied to real-world networks. Most of the social networks exhibit an assortative mixing pattern, where a famous person befriends another famous person, whereas technological networks (Internet) show a disassortative mixing pattern, where a client (having very little resource) connects to a server (with rich resources). Analysis of attacks on real-world p2p networks and their impact on the topology of the network is difficult as the interconnections among the peers are not random; rather they evolve based on the needs of the connected peers and this brings in degree-degree correlation in the network.

To address the aforementioned issue we have developed an analytical framework to analyze the change in topology of a correlated network and propose a generalized model based on percolation theory to measure the resilience of a correlated network against any arbitrary attack. We have also defined metric to measure the critical condition for stability using our model. We have shown that the framework can also be applied to random networks. We have validated our theoretical results on the real-world representative snapshots of commercial Gnutella networks and the results are in very good agreement. In order to grow deeper insights we have used our model to study the following:

(a) Dependence of percolation threshold of a superpeer network on its peer degree, superpeer degree at different levels of degree-degree correlation

(b) Impact of different attacks on the topology of a network at different degree-degree correlation.

(c) Impact of attacks on the degree-degree correlation of a network.

(d) Determining the parameters that govern the stability of superpeer networks

Anup Kumar Bhattacharya
Email: anup@cse.iitkgp.ernet.in
Joined the department in: September 2008

Anup Kumar *Bhattacharya received his B.E. degree in Electronics & Telecommunication Engineering from Jadavpur University, Kolkata in 2006. From July, 2006 till August, 2008, he worked in PricewaterhouseCoopers, Kolkata, as a consultant. Since September, 2008, he has been a research scholar in the department of Computer Science & Engineering in IIT Kharagpur. His research interests are in the areas of Algorithms & Cryptography.*

Supervisors: Prof. Dipanwita Roychaudhury and Prof. Abhijit Das

Algorithm Design and Implementation issues in cryptography

Efficient software implementation of cryptographic primitives: Pairing based cryptography is used nowadays to design different cryptographic protocols like signature generation and verification etc. In this work, we are concentrating on efficient implementation of pairing on super-singular elliptic curves defined over characteristic 2 & 3.

We design and implement efficient arithmetic for characteristic 2 & 3 fields. We implement variants of addition, multiplication, square, square root and inverse routines for these finite fields.

Using these routines, we implement a variant of pairing defined over super-singular curves named Eta pairing. In this work, we are also exploiting the architectural facilities like SIMD parallelism for faster implementation of pairing. We are comparing between horizontal and vertical parallelizations using SIMD intrinsics.

Binanda Sengupta
Email: binujucse3@gmail.com
Joined the department in: July 2010

Binanda Sengupta *received a B.E. degree in Computer Science and Engineering from Jadavpur University, Kolkata in 2007. From October 2007 till January 2010, he worked in Tata Consultancy Services as an Assistant Systems Engineer. Since July 2010, he has been a research scholar (MS) in the department of Computer Science & Engineering in IIT Kharagpur. His research interest area is Cryptography.*

Supervisor: Prof. Abhijit Das

Ciphertext-only Attacks

Cryptanalysis is the method to retrieve the key of an entity using the given cryptosystem. Ciphertext-only attack is to find the actual key based on some given ciphertexts only. Brute force search is one of the techniques that can be used in this purpose, but it leads to exponential running time. We have made a parallel implementation of exhaustive search when some plaintext-ciphertext pairs are given. We are trying to port it in ciphertext-only situations also. We are also trying to exploit some redundancy in plaintexts to reduce the search space substantially.

Biswajit Das
Email: bdas@cse.iitkgp.ernet.in
Joined the department in: July 2009

***Biswajit Das** received B.Tech. degree in Electronics and Communication Engineering from Murshidabad College of Engineering and Technology, Berhampore in 2006. From November, 2006 till January, 2008, I have worked in department of Mining Engineering, IIT Kharagpur, as a Junior Project Assistant. Since February, 2008, I am working as Junior Project Officer in the department of Computer Science & Engineering in IIT Kharagpur. Since July, 2009, I am pursuing MS in the department of Computer Science & Engineering in IIT Kharagpur. My research interest are in the areas of Speech Recognition and Speech Enhancement.*

Supervisor: Prof. Pabitra Mitra

Elderly Speech Recognition

Speech is the most commonly used and straightforward way of communication. Speech can be another interface to computer. Most of the work on speech related application has been done for younger user group. In our society, there is a significant percentage of older people. Older people need variety of spoken dialogue system which will be more entertaining to them. They are the last user group to take advantage from computer. Senior people mainly with arthritis and low vision can access the computer via their voice. There are many reasons behind not having used computer like physical difficulty to access mouse or keyboard, reluctance to use new technology. Main challenge is to build up Automatic Speech Recognition (ASR) system for older people which will give good performance.

With increment of age, several changes take place in our articulatory system. Our articulatory system consists of nose, mouth, pharynx, velum, epiglottis, vocal folds and lungs. With ageing, reduction in respiratory muscle strength, loss of elasticity stiffening of the thorax and loss in the diaphragm strength are the most significant changes in the respiratory system [1]. Due to this, amount of air comes out from lungs decreased for the older people.

Generally age related changes occurred in anatomy and physiology of the speech mechanism, decline sensory feedback, reduced speed accuracy of motor control, and diminished cognitive linguistic function. With ageing, changes in the larynx, such as stiffening of the cartilages to which the vocal cords are attached and degeneration of intrinsic muscles which reduce the ease of vocal fold adjustments during phonation. Increase in the stiffness of vocal cord cover is also observed, leading to instability of the vocal fold vibrations. Thickening of laryngeal epithelium progressively with age has been reported in many studies which may contribute to the lowering of fundamental frequency and increased harshness observed in older voices.

Changes in the vocal cavity consist of degeneration of pharyngeal muscles, reduced in salivary function, decreased tongue strength and tooth loss. Degenerative changes are also observed in the temporomandibular joint which controls the jaw movement during speech production. These changes could considerably affect the articulation of speech. Changes in vocal tract dimensions have also been observed in older speakers, which may affect the resonance patterns in older speakers resulting in reduction of articulatory precision.

These physiological changes make differences speech characteristics of older from younger people. Different type of variability induced in older speech signal like Fundamental frequency (F0), Formant frequencies (F1, F2, F3...etc), Jitter, Shimmer, Harmonic-to-Noise ratio (HNR) etc. It has been observed in different studies [2] that Word Error Rate (WER %) are more for older people. My interest of research is to build a model that will reduce the variability among speakers.

There is no speech corpus of elderly people in Indian languages. I have to develop a Bengali speech corpus of elderly people. After processing the speech signal, I have to create acoustic model and language model for elderly people. Main goal of my research is to build up a ASR for older people.

References

1. P. B. Mueller, “The aging voice”. Seminars in Speech and Language, 18, 159169, 1997.
2. P. Torre III, J.A. Barlow “Age-related changes in acoustic characteristics of adult speech”. Journal of Communication Disorders, 42 (5), pp. 324-333. 2009

Biswanath Barik
Email: bn.barik@gmail.com
Joined the department in: January 2008

Biswanath Barik *received B.Tech. degree in Information Technology from University Science Instrumentation Centre, University of Kalyani, Nadia, West Bengal in 2003. From July 2005 till July 2007, he worked in Mallabhum Institute of Technology, Bishnupur, Bankura, West Bengal, as a Lecturer in the Department of Information Technology. Since August 2007, he has been working as a Junior Project Officer in a DIT sponsored project. In January 2008, he has joined in MS (research) programme of Computer Science & Engineering Department of IIT Kharagpur. His research interests are in the areas of Natural Language Processing, Computational Linguistics and Machine Learning.*

Supervisor: Prof. Sudeshna Sarkar and Prof. Anupam Basu

Relevance of Bengali Chunking in Bengali to Hindi Machine Translation

Text chunking is the task of dividing a sentence into syntactically correlated non-overlapping group of words, called chunks. The meaning of word chunk differs from psycholinguistic theoretical explanation to computational perspective. In our work we consider chunk as "*a non-recursive phrase consisting of correlated, inseparable adjacent words governed by the head of the chunk*".

The chunking task can be viewed as a segmentation and classification problem. The input to a chunking system is a sentence (i.e., a sequence of words). The chunking system (or chunker) divides the sentence into a number of meaningful segments (or word chunks) based on the local dependencies among the adjacent words. The segments are then classified and labeled according to the property of the head of the chunk.

Chunk identification is a pre-requirement in many Natural Language Processing (NLP) tasks. Chunking is also useful in Information Retrieval (IR) and Text to speech (TTS) system. We have developed a statistical chunking system for Bengali. As Bengali is morphologically very rich language, a set of diverse linguistic features play important role in chunking. We have explored these linguistic features and identified the best feature set correspond to the best chunking model. We got a reasonably good accuracy chunking system.

We are currently studying the importance of chunking in Bengali to Hindi Machine Translation. Bengali and Hindi languages have very similar language structure i.e., S-O-V structure. They also share a substantial amount of lexicon. However, a word to word translation from Bengali to Hindi does not produce a good quality translation. On the other hand, the rule-based machine translation approaches require the complete syntactic analysis of the input sentence. To achieve such analysis, a sentence level parser is required. But, till date Bengali does not have a good quality parser. Interlingua machine translation, on the other hand, requires semantic representation as well as full syntactic analysis. In this scenario, we are investigating the appropriateness of chunk to chunk translation from Bengali to Hindi. We are limiting the source language sentence analysis up to chunking and translating each source language chunk to the target language. Finally, we are also exploring what type of translation errors occur during the translation process and how those errors can be removed so that good quality translation can be achieved.

Biswanath Saha
Email: itsmebiswa@gmail.com
Joined the department in: January 2010

Supervisors: Prof. A. K. Majumdar and Prof. Jayanta Mukhopadhyay

Baby Cry Analysis

I am working on Newborn Baby Cry Analysis, which is a part of the project PCS being funded by DIT, MCIT, Govt. of India. The work includes diagnosis of the reason of cry of the newly born babies who may or may not have medical sickness. Currently I am collecting the data from the Dept. of Neonatology, SSKM Hospital, Kolkata. An Automatic Baby Cry Recorder has been developed and installed for recording the cry of the babies in the Neonatal ICU of the hospital. The recorded data are then brought to our laboratory for further research analysis which includes some speech processing techniques. The ground truth data have been labeled by the doctors with different reasons of cry such as ‘Hunger’, ‘Wet Diaper’, ‘Sleepy’, ‘Hold The Baby’, ‘Too Hot/Cold’, ‘Pain’, ‘Cold’, ‘Annoyed’ etc. Based on some training data the objective is to correctly classify the new data in the appropriate category of the reason of cry.

Chandan Misra
Email: chandan.misra1@gmail.com
Joined the department in: January 2010

***Chandan Misra** received a B.Tech. degree in Information Technology from Kalyani Govt. Engineering College, Kalyani in 2007. From June 2007 till February 2009, he worked in Wipro Technologies, Bangalore and Chennai, as a Software Design Engineer. Since January 2010, he has been a research scholar in the department of Computer Science & Engineering in IIT Kharagpur. His research interests are in the areas of Music Computing, Natural Language Processing and Digital Signal Processing.*

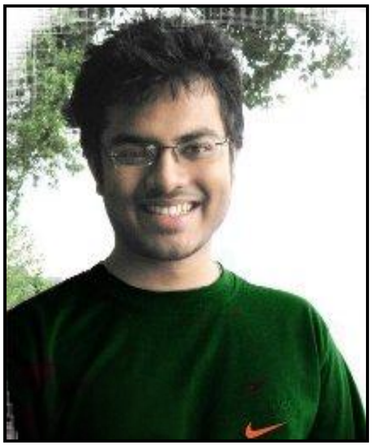
Supervisor: Prof. Anupam Basu

Transcription and generation of musical notes and Analysis and Classification of ornaments in North Indian (Hindustani) Classical Music

North Indian Classical Music also known as Hindustani Classical Music is one of the oldest music cultures still being performed actively. Although technology related to music analysis have taken a giant leap over the past few years, not much has been researched related to the transcription and expressiveness of various genres of Hindustani Classical Music.

In the current work, we have divided the problem in two parts. First to transcript it and then analyze it to generate melody as per the musical standard. We have taken several other musical genres and their notation systems for transcription purpose. There was no standard format to encode the musical symbols in computer. So, we have developed one system which contains all the musical symbols of various notation systems used by different genres. Moreover we have proposed to encode some symbols in the Unicode standard so that the system can be more unique.

The system will store all the musical information that is required to generate music from the music piece. We have used a customized Java API in order to produce musical sound from the computer sound card. But there are some ornaments present in these genres based on micro vibrations and changes in the pitch information. Currently we are working on the pitch analysis of various musical events.



Debmalya Sinha
Email: debmalya.01@gmail.com
Joined the department in: August 2009

***Debmalya Sinha** received B.Tech. degree in Computer Science from Aryabhatta Institute of Engineering and Management, Durgapur in 2009. Since August 2009, he has been a research scholar in the department of Computer Science & Engineering in IIT Kharagpur. His research interests are in the areas of Human Computer Interaction and Cognitive Science.*

Supervisor: Prof. Anupam Basu

Sahaj Linux

According to the common beliefs of the segment of non-technical computer users over the world, Linux is mainly a command line based operating system with an unfriendly Graphical Interface. The geeky commands, diversified settings and a number of advanced functionalities is firstly hard to learn for a beginner and secondly, are of little to no use for the common people like users in rural India. For the Windows users, the root directory oriented perspective of the Linux file-system is also a reason for users to stick with the vulnerable and costly but simple Windows operating system.

My focus is into the mentioned segment of common computer users and to develop an intuitive, easy-to-use interface which, ported to Ubuntu Linux, will be called "Sahaj Linux". The interface provides an intuitive and minimalist approach to common day-to-day tasks of a general computer user coupled with the flexibility and power of a Linux system.

The main focus is on a new approach for the File Browser, named SahajBrowser. The SahajBrowser is specially designed to resemble the common mental model of the directory tree structure among users for increased performance in terms of reduced number of clicks in regular use. The Unique feature of SahajBrowser is, the user can open and see multiple directories at one time. This makes File management and Browsing much easier.
My primary research goal is to make the interface better usable for the users in rural India. To understand what the users want and how to present it to them such that it minimizes the effort for learning while being more time efficient as well, we used a few Cognitive Models to predict the efficiency of using SahajBrowser compared to other existing Browsers like Windows Explorer in Windows, Nautilus, Konqueror in Linux and “Find” in Macintosh Systems.

Most of the Usability evaluation techniques till date either use only the system specification for determining good or bad interfaces, or they are entirely based on the questionnaire to the users. There are a lot of scopes for future work to be done with more precise usability evaluation with taking these two schools together.

I'm currently working on for constructing mathematical representations of the SahajBrowser system so that it is logically comparable to existing browsers. After the present work is done, this research has scope for extending usability evaluation using "syndesis" with various human emotions like moods, emotional states and Persuasion.

Partha De
Email: partha.de@cse.iitkgp.ernet.in
Joined the department in: July 2010

***Partha De** received a B.Tech degree in Computer Science from West Bengal University of Technology, Kolkata in 2005, and Post-graduate Diploma in Information Technology (PGDIT) from Indian Institute of Technology, Kharagpur in 2007. From September 2007 to June 2008, he worked, as a Program Facilitator. From July 2008 to July 2010, he worked in the "India Chip Design program" at Indian Institute Technology, Kharagpur as a Junior Project Assistant. From July 2010 he has been a research scholar in the department of Computer Science & Engineering in IIT Kharagpur. His research interests are in the areas of High-level Synthesis and low power VLSI design.*

Supervisor: Prof. Chittaranjan Mandal

Structure Architecture Driven High-level Synthesis for Array Intensive Applications

High-level synthesis (HLS) is the process of generating register transfer level (RTL) designs from a given behavioral description. To deal with the increasing complexity of today's VLSI designs, the use of HLS tools becomes increasingly crucial towards raising the level of abstraction in system design. Over the last several years, various such HLS tools have evolved producing elementary non-optimized data paths to more sophisticated ones generating data paths optimized for area, wire length, time, power, etc. Some of the existing HLS tools emphasis on optimization of layout area/wire length of the output RTL without considering an organization of the final data path at the start of the HLS process. We believe a better organization of the datapath and an abstract view of it at the input to a HLS tool along with the input behavior should aid the generation of optimized RTL with respect to layout area as well as wire length. With this objective, a HLS tool named “SAST” (Structured Architecture Synthesis Tool) has been developed in our group. A *simple but predictable* architecture called *structure architecture (SA)* has been proposed and datapaths are synthesized by SAST to conform to that structure. SAST takes a behavioral description written in C-like language along with the parameters of the SA and generates synthesizable RTL. The SA is organized as architectural blocks (A-blocks). Each A-block has a local functional unit, local storage. All the A-blocks in a design are interconnected by a number of global buses. So, the structure of the final architecture is fixed at the start of the synthesis but the final interconnection will be finalized during the synthesis procedure. The advantage of this architecture is that the user has the full control over the final architecture. Also, this structure data paths avoid random interconnects between data path components. The objective of my work is to validate our claim by extensive experimentation's. For this purpose, the RTLs generated by SAST from various benchmark problems need to be synthesized further with commonly used EDA tools such as Synopsis DA (for logic synthesis) and SoC Encounter from Cadence (for physical design) to obtain the actual measure of the layout area and wire length of the chip.

I am now evaluating SAST with respect to the final design produced by other behavioral synthesis tools. I am also evaluating the benefits of the SA methodology and exploring ways to improve designs generated by SAST. Also, SAST does not currently support global memory which is useful for supporting array intensive behaviors. My next objective is to extend SAST to support global memory. Initially, we plan to connect the memory to the global buses in SA. But, this organization will increase the number of transfers over global buses. Therefore, we need a better organization of the data path for array intensive programs. My third objective is to evolve the suitable architecture and subsequently tune the phases of HLS process accordingly for array intensive programs.

Praloy Kr. Biswas
Email: Praloy.biswas@cse.iitkgp.ernet.in
Joined the department in: August 2009

***Praloy Kr. Biswas** received his B.E. in Computer Science and Engineering from Jadavpur University, Kolkata in 2005. He started his professional career as a project engineer in Wipro Technologies, Hyderabad. Then he moved to EDA domain and served as Support Engineer in Atrenta India Pvt. Ltd, Noida and subsequently to Verific Design Automation (in Kolkata division) as software engineer, mostly involved in developing analyzer and RTL elaborators for HDLs (primarily for Verilog). Currently, he is pursuing his MS from IIT, Kharagpur in Cryptography and Network Security.*

Supervisor: Prof. Dipanwita Roy Chowdhury

Use of Algebraic Geometric Techniques in Cryptography and various other field of computational commutative algebra

In short, Algebraic Geometry focuses on how to tackle Geometric problems with Algebraic means. The key to the relationship between these two seemingly separated branches of Mathematics is the observation that a polynomial $f(x_1..., x_n)$ in $k[x_1,…,x_n]$ gives a function from k^n to k, where k is the underlying field. More to it, well studied algebraic structures like Ring, Ideal can be used while dealing with Geometric problems, if we introduce the concept of Variety of a system of multinomials, which is nothing but the solution set of the system. The set of multinomial having a fixed Variety form an Ideal and there is a well known correspondence between these two - i.e. between the Ideal and the corresponding Variety - the so called "Ideal-Variety correspondence Theorem", due to David Hilbert.

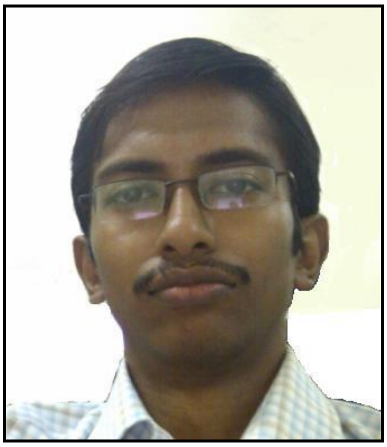
Among the bases that generate a given polynomial Ideal, there are few having a special property, which helps a great deal to solve many Geometric as well as Algebraic problems computationally. These bases are called Grobner Basis. In 1965, Bruno Buchberger, devised an algorithm which can calculate such basis for a polynomial Ideal and thus created a huge scope to tackle many Algebraic and Geometric problems computationally.

In our work, we developed a tool which computes the Grobner Basis for a system of multinomials and thereby finds out the solution set for the system. And in so doing we had too developed a class comprising of the operators which can do multi-precision arithmetic. This is for Rational field, where we can further apply this tool in many upcoming areas like Algebraic Statistics, Automatic Geometric Theorem proving etc.

The other part of the work which solely aims at using these techniques to algebraically crypt-analyze the stream ciphers like Bivium or Trivium, is underway too. There we have developed the arithmetic in GF (2) and currently trying to handle large amount of equations, which we have to generate in order to crypt-analyze the ciphers mentioned above.

References

1. Buchberger "An algorithmic method in polynomial ideal theory. in Multidimensional systems theory", ed. by N.K. Bose. D Reidel Publishing Company, Dordrecht, (1985), 184–232.
2. Cox Q., Little J, O'Shea D. "Ideals, Varieties, and Algorithms", Springer Verlag UTM.
3. Gregory V. Bard, "Algebraic Cryptanalysis", Springer.
4. Christophe De Canniere and Bart Preneel, "TRIVIUM specifications", Project estream, URL: http://www.ecrypt.eu.org/stream/p3ciphers/trivium.



Pramit Roy
Email: pramit@pramitroy.com
Joined in the department: September, 2008

***Pramit Roy** received a B.E. degree in Computer Science and Engineering from Jadavpur University, Kolkata in 2008. Since September 2008, he has been a research scholar (MS student) in the department of Computer Science & Engineering in IIT Kharagpur. His research interests are in the areas of Vehicular Ad Hoc Network. .*

Supervisor: Prof Arobinda Gupta

Misbehavior Detection in VANET using Secondary Information

A mobile Ad Hoc Network set up in vehicles is called the vehicular Ad Hoc Network. There are certain applications in VANET where vehicles raise alert messages for some safety related events like crash, slow speed or stopped motion, road hazard, emergency braking etc. Vehicles send the alert from an on-board unit or the OBU which can be manipulated to raise false alerts. We call act of sending false or wrong alert messages as misbehavior and the vehicle involved in it to be misbehaving vehicle. There exist certain authentication methods to abstain a recognized (to be misbehaving) vehicle from communicating with other vehicles, but no sufficient methods to detect a misbehaving vehicle. In VANET there are certain safety events which can occur as a consequence of another event, like emergency braking or slowing down of vehicles for a crash which has happened ahead. In my work I propose a misbehavior detection scheme using alert messages (secondary information) which are raised as a consequence of the alert message (primary information) to be detected as true or false. A vehicle raises a PCN alert when it crashes and SVA alert when it stops or moves with an unusually low speed. My work specifically encompasses the detection of the truth of PCN alert by using SVA alerts as secondary information.

Rajdeep Mukherjee
Email: rajdeep.mukherjee@cse.iitkgp.ernet.in
Joined the department in: July 2010

Rajdeep Mukherjee *received a B.Sc. degree in Computer Science from Asutosh College, Kolkata in 2007, and B.Tech. degree in Computer Science and Engineering from University of Calcutta, Kolkata in 2010.Since July 2010, he has been a research Consultant and MS student in the department of Computer Science & Engineering in IIT Kharagpur. His research interests are in the areas of Formal Verification, Low-power Circuits and System, Low-power aware High-level Synthesis, Fine grained Power Management.*

Supervisors: Prof. Pallab Dasgupta and Prof. Ajit Pal

Low-Power Aware High-level Synthesis Using Fine-Grained Dynamic Voltage Scaling

Power consumption has become an important factor in electronic portable system design, where excess power dissipation may lead to less reliability. Also, the demand for personal computing devices and mobile communication equipment is increasing. More and more consumers use small mobile devices in their daily life .The rapid progress in semiconductor technology has also led to higher chip density and operating frequency, which makes these mobile devices more complex and power-consuming. Therefore, power consumption has become an important issue in circuit design for this mobile devices.Advances in silicon technology also enables entire systems to be integrated on a single chip, which is known as Systems-on-a-Chip (SoC). The development of low power devices, circuits, algorithms, architectures, and CAD tools are fundamental for the successful realization of SoC.

The work focuses on the power optimization (minimize average and peak power) at the scheduling stage of high-level synthesis. Scheduling in high-level synthesis stage can be Latency- Constrained , Resource- Constrained Scheduling and latency and resource constrained scheduling.Transformational algorithms carried out in the scheduling stage of High-level Synthesis are used to obtain a new schedule from a default schedule by applying transformations. The fundamental transformations consist of moving operations or blocks executed serially into the same execution interval (parallel execution), or moving operations executed concurrently into subsequent time steps (serial execution). These approaches use Branch and Bound algorithms, ILP formulations, Trace and Percolation Algorithms, and Stochastic techniques.

The average power dissipation P of a digital CMOS circuit is composed of two components:

P = Pstatic + Pdynamic

Static power consumption can be described as the following equation:

Pstatic = I leakage × V

where I leakage is the leakage current, and V is the operating voltage.

Pdynamic is the dominant part of the power dissipation in CMOS circuits, which can be decomposed into the following three terms:

Pdynamic = P switching + P short-circuit + P leakage

Power savings can be achieved by reducing the following parameters:

- Switching activity
- Load capacitance
- Supply voltage
- Clock frequency

The research mainly concentrates in reducing power consumption of functional units that operates at multiple-voltages (5 V, 3.3 V, 2.4 V, and 1.5 V). This is mainly achieved by dynamically scaling supply voltages or threshold voltages of the functional units available during the allocation stage of High-level synthesis to reduce the dynamic power and leakage power respectively. A Low-Vdd resource operates slowly thus consuming more time steps compared to resources that operates at high-Vdd. Contrary to this, a functional unit that operates on high-threshold voltage is slower compared to the one that operates on a lower threshold volatge. Hence , the critical path of the Data Flow graph must contain the faster units , that is one with high supply voltage and low-threshold voltage and the non-critical path must utilize the slack by placing the slower functional units thus reducing the overall power consumption. Further, the idea to dynamically scale the resources to operate at different voltages at different instance of time has lowered the total number of resources that would otherwise be required in Multi-voltage scheme. This technique of managing power is also referred to as Fine-grained Power Management. There has also been research done in low power design at the behavioral level, such as reducing the number of registers and switching activities in registers, as well as efficient register allocations.

Ramji Nagariya
Email: ramjinagariya@gmail.com
Joined the department in: July 2010

***Ramji Nagariya** has completed his B.tech (Computer Science & Engineering) from IIT Kharagpur in 2007. He has joined the MS Program in the same department since July, 2010. His research interest is in area of complex networks and its applications. He also has strong inclination towards social problems facing the nation and has been involved in social activities.*

Supervisor: Prof. Niloy Ganguly

Analysis of One Mode Projections of Bipartite Networks

We make an attempt to understand the properties of one mode projection of a growing bipartite network when one set of node is kept fixed and the other set is growing. The edge weights between two nodes in the one mode projection on fixed set is given according to number of nodes in growing set which are connected to those two nodes in fixed set. Any new node entering the growing set attaches to a few nodes in fixed set according to preferential attachment. Given a fixed set of size m, we try to investigate at what value of size n of growing set the one mode projection on fixed set becomes a clique.

Ritwika Ghose
Email: ritwika.ghose@gmail.com
Joined the department in: September 2009

***Ritwika Ghose** received B.E. degree in Information Technology from West Bengal University of Technology, Kolkata in 2009. Since September 2009, she has been a research scholar in the department of Computer Science & Engineering in IIT Kharagpur. Her research interests are in the areas of Natural Interface Design, Assistive Technology and Human Computer Interaction.*

Supervisor: Prof. Anupam Basu

Open Source Web Browser for Visually Handicapped People

With the advancement of World Wide Web, the popularity of using Internet increased manifold. People could now access huge volumes and variety of information very easily. Apart from this, a lot of other factors like communication, and easy access to different types of services like banking, ticket reservation etc. also led to the widespread use of WWW. The importance of Internet leads to another important associated factor that is the interface which brings forth the information to the end user. Over the years, a number of web browsers have been designed based on the requirements of normal users. Besides providing advanced functionality, the developers of web browsers have recently started concentrating on factors like easy access to the various possible operations, fast learning of the usage and functions of the interface, ease of use, etc. Human Computer Interaction and User Centered Design started gaining popularity. A number of theories emerged which provided guidelines for designing user interfaces. Ben Shneiderman provides eight golden rules [1] for designing a good interface focusing on the user and his ease of using the interface. Another set of principles suggested by Larry Constantine and Lucy Lockwood [2] emphasized that it is not sufficient to just focus on users, but more specifically, it is important to focus on the usage i.e. their work and providing usable tools for them.

Most of these principles, however, concentrated on increasing usability for people without any kind of disability. However, there is a significant population of users, who are challenged in different aspects of physical and/or mental abilities. Along with problems like cognitive and neurological disabilities, aging-related conditions, motor disabilities, a large number of people are affected by visual impairment. The interface that has been designed for a sighted user would not directly work for a blind person because most of the design elements which aimed to increase usability for a sighted person are based on visual aids. For example, an important design style of interaction suggested in the theories of Human Computer Interaction is WIMP which stands for “window icon menu, pointing device”, which is the basis of design of the existing web browsers. The users are presented with a window, which has a number of easily accessible icons corresponding to some operations. Moreover, menus are provided to categorize the operations under a single heading and are brought in view only when they are clicked. It is easy for a user to locate a desired operation by quickly scanning through the menus. Further, the selection of a menu item, are often accompanied by change of color of that particular area, so that the user has a clear idea where his mouse is pointing to. When this same design is mapped to a visually handicapped person, the whole scenario is changed. All the factors of the interface, i.e. window, menu, icon color, etc. have to be represented in a single dimension, i.e. speech. The main purpose of icons i.e. to make an option easily available is lost when it has to be presented in speech. If a non-sighted user wants to search an operation in a menu through a speech based interface, he has to hear through all the operations linearly until the desired one is reached. This is a much more time consuming process than normal visual scanning. Moreover, unlike a sighted user who can directly move to the location of the menus to activate them, a blind user has to recall the methods (key combinations to be pressed) in order to activate a menu, due to which the users’ memory load automatically increases. So it is necessary that the interface designed for a visually handicapped person has different features modes of presentation to enhance usability.

The other important factor of web browsing is to present the web page content to the user. A very important challenge in this aspect is to give a blind user similar experience and comforts in web browsing as a sighted person. Navigation and keeping track of links, skipping content to go to a desired location in the page, quick scanning to have an idea about the content of a web page are some of the major challenges that are required to be handled when presenting a web page to a non-sighted person.

Based on the above perspective, the goal of the present research is to develop a browser for the sightless people, so that they can browse the web. Specifically, the objective of the research is to identify the cognitive needs of the blind users and provide interface features through non-visual means enabling them to browse. The present research will explore speech interface, gesture based interface, Braille interface and speech feedback enabled keyboard interface as possible input-output means. Another aspect to be studied is the plan of navigation through the interface that will be better suited for the blind users. Such plan design will also necessitate categorization of the different control and access points, to minimize cognitive load of the sightless users. The latter will also depend on the category of the users (novice, semi-expert, expert etc.) and hence the plan should either be adaptive or separate plans should be adopted for the different categories. The other aspect of the research focuses on the representation of web page content in a hierarchical fashion in order to maximize comprehensibility and easy navigation through links. Along with this, field testing for validation will also be done at various stages.

References

1. Ben Shneiderman, Designing the user interface: strategies for effective human-computer interaction, Addison-Wesley Longman Publishing Co., Inc., Boston, MA, 1986
2. Constantine L.L., and Lockwood L.A.D. (1999) Software for Use: A Practical Guide to the Models and Methods of Usage-Centered Design. ACM Press, New York.

Sandipan Mandal
Email: mandal.sandipan@gmail.com
Joined the department in: August 2008

***Sandipan Mandal** received B.Tech. degree in Computer Engineering from Malviya National Institute of Technology, Jaipur in 2005, and now pursuing M.S. in Computer Science and Engineering in Indian Institute of Technology, Kharagpur . From November 2005 to May 2008, he worked with Tata Teleservises Ltd., as a Senior Engineer. He worked as Junior Project officer in Department of Computer Science and Engineering from August 2008 to July 2009. Since July 2009, he is an M.S. scholar in the department of Computer Science & Engineering in IIT Kharagpur. His research interests are in the areas of Speech Recognition and Assistive Technology*

Supervisor: Prof. Pabita Mitra

Bengali Speech Recognition and an Application for Visually Impaired Community

Speech Recognition is a process for converting spoken words or sentences to the corresponding textual representation. This has a wide domain of applications Dictation system, Voice Command-based application e.g., voice dialing, Data entry and many more in Human Computer Interaction (HCI) field.

Although Bengali is currently one of the most widely spoken language in the world, there has been relatively little speech recognition research on Bengali compared to the other languages. In order to develop a Bengali ASR System, we use an open source speech recognition engine Sphinx-3 for both training and test case. Sphinx-3 is HMM-based, speaker-independent, continuous recognition system, capable of handling large vocabularies. To develop Bengali ASR along with HMM model we use another statistical model: Trigram Language Model. For state probability distribution it uses continuous density of Gaussian Mixture distributions. All phonemes are modeled as a sequence of HMM state and likelihoods (emission probability) of a certain frame observation is produced by using traditional Gaussian Mixture Model (GMM). Different parameters in the system were adjusted, some are training parameters (number of state per HMM, number of Gaussians densities) and other are decoding parameters (silence insertion probability, word insertion probability, language weight). Best results were found at 3-state left-to-right architecture which is adopted to model each speech unit and each state was modeled using a mixture of 16 Gaussians.

Developing Speech corpus, corresponding Text corpus, Phoneme selection and Bengali pronunciation dictionary is part of this.

A SMS and an E-mail sending application has been developed for blind user by using speech recognition. Instead of writing blind person can give speech as input. Speech input is converted to corresponding text and this converted text is given to SMS and E-mail.

Satrajit Ghosh
Email: satrajit@cse.iitkgp.ernet.in
Joined in the department in: June 2008

***Satrajit Ghosh** received his B.Sc (Honours) degree in Physics from Scottish Church College, University of Calcutta in 2005 and B.Tech degree in Computer Science and Engineering from Department of Computer Science and Engineering, University of Calcutta in 2008. Since June 2008, he has been a research scholar in the Department of Computer Science and Engineering in IIT Kharagpur. His research interests are in the areas of Algorithm Design, Efficient and Parallel implementations and Cryptography.*

Supervisor: Prof. Abhijit Das

Algorithm Design and Implementation issues in cryptography

Algebraic attack on Block Ciphers: Block ciphers are an important building block of modern cryptography. In August 2000, the block cipher Rijndael was selected as Advanced Encryption Standard (AES). Rijndael is a key-iterated block cipher with a very strong algebraic structure. AES can be represented as algebraically closed equations over GF (28) or as a system of multivariate equations over GF (2) with plain-text, cipher-text and key bits as variables. The system of equations then can be solved for the key values given a few known plain-text/cipher-text pairs for a specific key. We have proposed a new heuristic to reduce the number of linearized equations for the XL (eXtended Linearization) method. We have tried our algorithm on small random sparse systems, and observed significant improvement in the performance of the XL algorithm. Our plan is to mount attack on AES like ciphers using our algorithm.

Sirsendu Mohanta
Email: sirsendu@cse.iitkgp.ernet.in
Joined the department in: July 2009

***Sirsendu Mohanta** received his B-Tech degree in Information Technology from the Kalyani Government Engineering College, Kalyani in 2008. He joined the department of Computer Science &Engineering at IIT Kharagpur in July 2009 as a project fellow for a project sponsored by Bhabha Atomic Research Centre (BARC), Mumbai. He is also pursuing his MS. (by research) since July 2009. His research interests are in the areas of software reliability and software engineering.*

Supervisor: Prof. Rajib Mall

Prediction of Software Reliability at the Early Stages of Product Development Cycle

Safety-critical applications, such as aviation, nuclear power generation, satellite communication are required to be highly reliable as failure of these systems may cause injury or death to human beings. Apart from the safety-critical systems, software has become the integral part of most of complex applications. Thus it is very important to ensure that the underlying software will operate correctly, perform its intended functions properly and deliver its desirable output. Reliability of a software is defined in terms of probability of failure-free software operation for a specified period of time in a specified environment. Several approaches that have been reported in the literature to quantify the software reliability can be classified into two categories: reliability estimation approaches and reliability prediction approaches.

Reliability estimation approaches capture the failure behaviour of a program during testing phase and fit the failure data to a reliability growth model to quantify the reliability of the software. The main weakness of reliability estimation approaches is these approaches collect the failure data during testing or maintenance phases and estimate the reliability of the software. Therefore, if it estimated that a product would fall short of its required reliability goals, it becomes too costly to rework design or code to improve its reliability. Therefore, prediction of software reliability early in the product development phase is important for minimizing the rework costs.

Existing reliability approaches which predict reliability of the software in the early stages of product development cycle take software architecture into consideration. However, most of these approaches are based on some assumptions that make the reliability estimation too optimistic relative to real situations. First of all, these software architecture-based reliability approaches assume that system is composed of several components whose reliabilities are known. Secondly, components are assumed to be independent of each other. Usually, reliability figures of the components are not available and the components of the system are directly or indirectly influence each other through control and data flows.

In light of the above discussed inadequacies of existing approaches, we propose a bottom-up software reliability prediction approach that we named as Early Software Reliability Assessment (ESRA). Our ESRA approach focuses on predicting the reliability of object-oriented programs at the early stages of product development. In this approach we consider classes as the basic components and predict their reliabilities first. However, accurate reliability prediction of classes early in the product development cycle is a major challenge, as failure information neither from field nor from testing is available. In this context, we construct a fault model to categorize different kinds of possible faults and identify the different design metrics that are highly correlated to different categories of faults. We construct a Bayesian Belief Network (BBN) to determine the likelihood of the code faults from the identified design metrics. Based on the analysis of likelihood of code faults reliability of a class is predicted. We predict the use case reliabilities by using predicted class reliabilities. The system reliability is predicted based on the predicted use case reliabilities and execution frequency of each use case.

Our proposed ESRA approach addresses two unrealistic assumptions that are used in software architecture-based reliability prediction approaches. The first assumption that reliability of the components is known is resolved in ERSA approach as the reliability of the components or the classes are predicted from design metrics. Secondly, at the time of use case reliability prediction, we consider error propagation among the classes that participate in the use case. Therefore, basic components or the classes are not at all considered as independent in ESRA approach.

From our empirical studies, we observed that our ESRA approach has gained more accuracy in reliability prediction compared to other approaches. We are currently carrying out more exhaustive studies over prototype implementations that represent few real world problems to evaluate the effectiveness of our approach.

Sourya Bhattacharyya
Email: sourya.bhatta@cse.iitkgp.ernet.in
Joined the department in: July 2009

I am a M.S. student in Computer Science and Engineering Department, Indian Institute of Technology, Kharagpur. I am working under supervision of Prof. Arun Kumar Majumdar and Prof. JayantaMukhopadhyay. Currently my research is based on detection of epilepsy and sleep wave cycle patterns in neonatal EEG signal. I have graduated from Computer Science and Engineering Department, Jadavpur University, in 2006. Then I worked as a design engineer in STMicroelectronics Private Limited. My work there was in multimedia domain, especially in video transcoding and bit rate control.

Supervisors: Prof. Arun Kumar Majumdar and Prof. Jayanta Mukhopadhyay

Neonatal EEG Signal Processing

Electroencephalogram (EEG) monitors cerebral electrical activities through electrodes placed on scalp and provides a sensitive real time graphical representation of brain function. Especially for neonates, neurophysiologic disorders and seizures are mostly diagnosed by visual inspection of EEG signals because neonates commonly do not exhibit clinical sign and symptoms for seizures.

EEG waves are classified according to their frequency as - β (> 13Hz), α (8 - 13 Hz), θ (5 - 7 Hz), and δ (0 - 4 Hz). Usually, types of activities seen in neonatal EEG are in the frequency ranges of δ and θ. α frequency component are rather infrequent. β activities, especially in high β frequency range (> 17 Hz) occur rarely.

Most common EEG abnormal pattern for neonate is burst suppression pattern. It is a pattern of high amplitude activity interrupted by relatively low amplitude activity (typically less than 20 µV peak-to-peak). Repeated occurrence of burst-suppression patterns is a sign of epilepsy.

Visual inspection and finding burst suppression pattern over a long recorded raw EEG signal is very time consuming. So, our aim is to provide automated detection of EEG patterns of interest so that doctors or technicians can quickly locate and view those suspected regions. Two approaches are there for automatic summarization of the EEG information. They are, 1) Amplitude EEG based detection, and 2) Raw EEG data processing and finding suspected regions.

Amplitude integrated EEG (aEEG) is basically a band pass filtered (filtering frequency range is 1.6 - 15 Hz), rectified, semi logarithmic compressed display of total recorded EEG. It is very helpful in detecting long term EEG trends (of several hours, for example). In practice, technicians or doctors can see the long term EEG trend for current neonate and visually identify the suspected regions from aEEG. Prolonged burst patterns can be identified through visual inspection of single channel or multiple channel aEEG patterns. However, short term bursts or patterns cannot be easily identified. So processing raw EEG data for detailed identification is crucial.

Identifying burst patterns (voltage of 75 - 250 µV for minimum 1 sec) from raw EEG requires modeling of background continuous EEG (voltage of 20 – 50 µV in general) and locating the bursts or high energy components within it. Suppression patterns are continuous low voltage (generally < 5 µV for minimum 2 sec) occurrence, which, if follows burst patterns, confirm presence of seizures.

In detection of high energy patterns within background EEG, non cerebral activities, termed as artifacts, might be detected simultaneously. Artifacts are sudden spike or abnormal waveform, easily distinguishable from background EEG. Physiological artifacts are generated from patient’s body other than brain, such as from muscle twitch, body movement, cardiac activity, respiration, eye blinking etc. Extra physiologic artifacts are generated outside body, like from equipments or environment, such as from loose electrodes, power line frequencies, external movements etc.

A burst pattern is a slow wave, usually of frequency within δ or θ range, intermixed with spike waves. Modeling of such a wave and correlation with detected wave pattern is essential to confirm presence of burst suppression pattern.

EEG of neonates is unstable, in a broad sense. Monitoring sleep states of neonate is essential for correct interpretation of EEG. To monitor the sleep wave cycle (SWC) regularity, we first target detecting presence of sleep spindles, which are rhythmic activities observed in stage 2 of non rapid eye movement (NREM) sleep EEG. Sleep spindles consist in sinus like bursts that increase and decrease progressively in amplitude, with minimum duration of 0.5 sec, and frequency within 12 - 14 Hz.

Current research is going in two parts. One is to generate the mathematical model of burst suppression and sleep spindle patterns so as to automatically reject other high energy non burst or artifact components. Another is to observe and model SWC regularity or non-regularity in normal and epileptic patients over long term EEG recording.

Sumit Das
Email: sumitdas@cse.iitkgp.ernet.in
Joined the department in: July 2008

***Sumit Das** received B.E. degree in Computer Science & Engineering from Jadavpur University, Kolkata in 2006. From August 2006 till January 2008, he worked in Tata Consultancy Services (TCS), Kolkata, as an Assistant Systems Engineer. Currently, he is an MS student in the department of Computer Science & Engineering in IIT Kharagpur. His research interests are in the areas of Natural Language Generation and Text Mining.*

Supervisors: Prof. Anupam Basu and Prof. Sudeshna Sarkar

Improving the Fluency of Bengali Text Generated by a Natural Language Generation System

Natural Language Generation (NLG) research aims at systems that produce coherent natural language text from an underlying representation of knowledge. NLG systems should produce text which (i) faithfully represents the relevant knowledge and (ii) sounds natural. These two goals of text generation are termed as fidelity and fluency. In natural language text, noun phrases (NPs) with multiple modifiers occupy a significant portion. Wrongly ordered premodifiers in such NPs affect the meaning and fluency of the text. Thus, correct ordering of the prenominal modifiers is an important task in text generation. On the other hand, unnecessary repetition of words makes text less fluent and non-coherent. In NLG, the task of combining simple text spans by removing the unnecessary repetitions is called text aggregation. Syntactic aggregation is the most prevalent form of text aggregation observed in real discourse. In syntactic aggregation the simple text spans are combined using linguistic rules resulting in more fluent, concise and coherent text. Thus, prenominal modifier ordering and syntactic text aggregation are two useful methods for improving the fluency of the text generated by an NLG system.

In the current work, we have the following objectives to improve the text fluency of Bengali text by an NLG system:

- To propose a method to correctly order the multiple premodifiers in an NP. We make an assumption that in an NP with more than one adjectival modifier, all of them modify the head noun. This assumption is valid because an NP with multiple modifiers, all modifying the head noun, is far more common than an NP with multiple modifiers where one of them modifiers another.

- To propose a method for performing syntactic text aggregation of simple text spans. In some cases rather than deleting the repeating entities, anaphoric pronoun generation improves the fluency of the generated text. Thus in addition to syntactic aggregation, our objective is to propose a method for generating unambiguous anaphoric pronouns in appropriate places.

To order multiple prenominal modifiers in an NP first we build a machine learning based model for pairwise modifier ordering. To order the modifiers of the seen modifier pairs we use direct corpus evidences from the training corpus. On the contrary, for ordering the modifiers of the unseen pairs transitivity and semantic clustering based methods are used. Using the learned pairwise orders the correct order of more than two modifiers is generated by following a graph based method.

In syntactic aggregation simple text spans are combined using linguistic rules. Thus, NLG systems should have the knowledge of the target language to perform this operation. We studied a corpus of Bengali sentences to identify the prevalent syntactic aggregation constructs in Bengali. We proposed rule-based approach for syntactically aggregating two simple sentences using the syntactic aggregation constructs identified. The inputs to this method are two simple sentences, the rhetorical relation connecting them and the appropriate discourse marker realizing the relation.

Currently, we are working on proposing a rule-based method for generating unambiguous anaphoric pronouns for repeating entities in text which cannot be deleted by syntactic aggregation.

Suprabhat Das
Email: suprabhat@cse.iitkgp.ernet.in
Joined the department in: July 2007

Suprabhat Das *had completed his B.Tech. degree in Computer Science and Engineering from Kalyani Government Engineering College in 2008. Currently he is pursuing his Master of Science (by Research) from Department of Computer Science and Engineering, IIT Kharagpur under the guidance of* Prof. PabitraMitra. *His research areas are Information Retrieval, Machine Learning and Stylistic Authorship.*

Supervisor: Prof. Pabitra Mitra

Information extraction and capturing stylistic attributes of a text document from Bengali literature

Indexing of a huge amount of documents and retrieval of data from that collection is a challenging task, especially for highly rich languages like Bengali. ‘Anwesan’ is an attempt to search from complete Rabindra Rachanabali collection. The search engine can search full-text data as well as metadata from the complete data set. To improve the search result and to retrieve relevant documents only, we have to extract root words from inflectional words while indexing.

Stemming is the process for reducing inflectional or derived words to its stem or root form. The process of stemming can not only enhance the recall but also reduces the index size drastically, particularly for morphologically rich languages like Bengali. A rule-based approach of stemming have been implemented to strip off the suffix part from the Bengali words and to extract the desired root word. This can be used with a renowned open-source text search engine, named Lucene. The stemmer has been tested on Bengali collection of the FIRE 2010 data set with 50 queries using Lucene as the search engine and it gives 96.27% recall and MAP value is equal to 0.4748.

Stylometry is the study of the unique linguistic styles and writing behaviors of individuals. Author identification is one of the important problems in stylometrics. Stylometry has been studied on English for long time. An initial attempt has been made to identify the authorship of a document from Bengali literary works. The writings of Rabindranath Tagore, Bankim Chandra Chattopadhyay and SukantaBhattacharyay have been taken into account for the study. Simple unigram and bi-gram features along with vocabulary richness were taken as feature set to discriminate amongst these authors. A classification accuracy of above 90% for unigram feature and almost 100% for bi-gram feature was achieved. Using some advanced feature or combination of many features on Bengali test collection may result better.

Many authors use to change their writing styles with time. A study on the change in writing style by Rabindranath Tagore with age is done. The complete collection of Rabindra Rachanabali has been divided into groups of two years interval. Distribution of total number of tokens and number of unique tokens with these two years of time intervals has been studied. From this study, it is obvious that Rabindranath Tagore also had changed his writing styles many times throughout his whole lifetime. This research work has facilitated to indicate where in the author’s career an undated text may fit.

Swarnendu Biswas
Email: swarnendu@cse.iitkgp.ernet.in
Joined the department in: September 2008

Swarnendu Biswas *received his B.E. degree in Computer Science and Engineering from the National Institute of Technology, Durgapur in 2005. He has worked as a software developer with Wipro Technologies from August 2005 to August 2008. He joined the department of Computer Science & Engineering at IIT Kharagpur in September 2008 as a project fellow for a project sponsored by General Motors India. He is also pursuing his MS. (by research) since January 2009. His research interests are in the areas of software engineering, real-time systems and embedded systems.*

Supervisor: Prof. Rajib Mall

Selecting and Optimizing Regression Test Suites for Embedded Programs

Regression testing is a popular but expensive technique to ensure the correctness of a program after it has been modified. In fact, regression testing has been estimated to account for almost half of the total software maintenance costs. Minimization of regression test effort is considered to be an issue of considerable practical importance, and has the potential to substantially reduce the software maintenance costs.

Regression test selection is one way of minimizing the regression test effort. Regression test selection techniques select a subset of valid test cases from an initial test suite to test the affected but unmodified parts of a program. A large number of regression test selection techniques for procedural, object-oriented, component-based, database, and web applications have been proposed in the literature. However, research results on regression test selection for embedded programs have scarcely been reported in the literature. Existing techniques may not work satisfactorily when used to select regression test cases for embedded programs. This is because these techniques ignore many important embedded program semantics during selection of test cases, e.g., tasks, timers, inter-task communication, etc. Regression testing of embedded programs is also constrained due to restrictions on the availability of resources such as time, budget, personnel, etc. In this context, paucity of time is considered to be the primary obstacle faced by testers during regression testing. Therefore, the set of selected regression test cases may be prohibitively large to be executed with the program under test and still meet the given constraints on regression testing.
In light of the above discussed inadequacies of existing approaches, we have proposed an improved regression testing technique for embedded programs. We first propose an intermediate representation

for modeling embedded programs developed in C language by extending the System Dependence Graph. Our proposed model in addition to capturing data and control dependencies, also represents important information such as, control flow, and embedded program features such as tasks, task precedences and inter-task communication using message queues and semaphores, etc. We have proposed a regression test selection technique based on an analysis of the constructed models. We select relevant test cases based on data and control dependencies by slicing the constructed models. In addition to data and control dependencies, our technique also takes into account the timing dependencies that arise among tasks in an embedded program due to precedence order relations, task priorities, inter-task communications and interrupt handlers.

We have also proposed a multi-objective regression test suite optimization technique using genetic algorithms. Test cases which execute critical functionalities and tasks whose timing behavior may potentially be affected are not omitted during our optimization technique. Additionally, our technique also aims to minimize the cost of regression testing, maximize the reliability of the frequently-executed functionalities and remove redundant test cases.

From our empirical studies, we observed that our regression test case selection and optimization techniques include all potentially fault-revealing test cases and at the same time achieve savings in terms of regression test effort without compromising on the thoroughness of testing. We are currently carrying out more exhaustive studies with industry example programs to evaluate the effectiveness of our approach.

Our Mentors:
Faculty of the Department

Jayanta Mukhopadhyay
Email: jay@cse.iitkgp.ernet.in

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).

Arun Kumar Majumdar
Email: akmj@cse.iitkgp.ernet.in

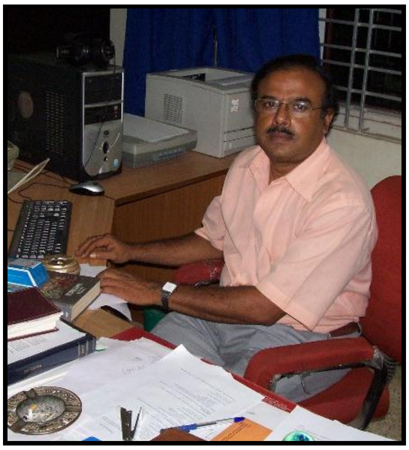
Research Interests: *Data and Knowledge-based Systems, Multimedia Systems, Medical Informatics, VLSI Design Automation*

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 again from 1998 to May 2001 and Head of Computer and Informatics Center, IIT Kharagpur: from 1998 to 2002

Arobinda Gupta
Email: agupta@cse.iitkgp.ernet.in

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 a 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.



Anupam Basu
Email: anupam@cse.iitkgp.ernet.in

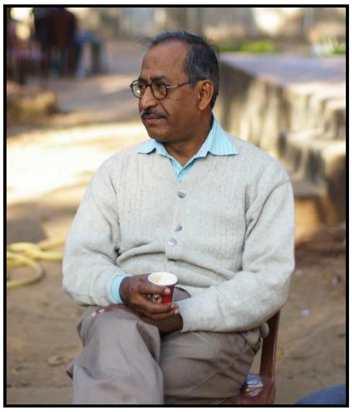
Research Interests: *Cognitive Science and Language Processing with particular focus on Intelligent Interface Design and Human Computer Interaction*

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.



Ajit Pal
Email: apal@cse.iitkgp.ernet.in

Research interest: *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.

Abhijit Das

Email: abhij@cse.iitkgp.ernet.in

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.

Chittaranjan Mandal

Email: chitta@cse.iitkgp.ernet.in

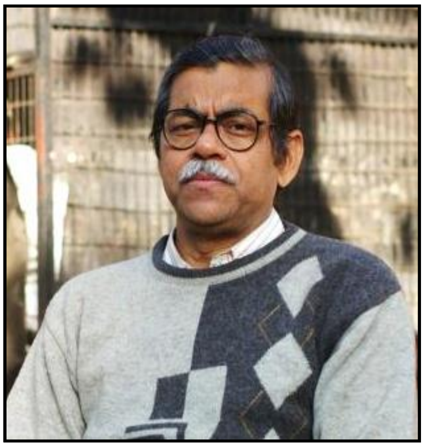
Research Interests: *Formal modelling and verification, high-level design, network and web technologies*

Chittaranjan Mandal received his Ph.D. degree from IIT, Kharagpur, India, in 1997. He is currently a 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 slso 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.

Debdeep Mukhopadhyay
Email: debdeep@cse.iitkgp.ernet.in

Research Interest: *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.



Dipankar Sarkar
Email: ds@cse.iitkgp.ernet.in

Research interest: *Formal Verification and Symbolic Reasoning*

D. Sarkar did his B.Tech., M.Tech. in Eletronics 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.

Dipanwita Roy Chowdhury

Email: drc@cse.iitkgp.ernet.in

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

Dipanwita Roy Chowdhury is a Professor in the Department of Computer Scienece 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 Pricipal Investigator of several R&D projects. She is the recipient of INSA Young Scientist Award and Associate of Indian Academy of Science. She is a fellow of the Indian National Academy of Engineering (INAE).

Goutam Biswas

Email: goutam@cse.iitkgp.ernet.in

Research Interest: *Theoretical computer science, compiler*

Indranil Sengupta
Email: isg@cse.iitkgp.ernet.in

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. He served as Head of the Computer Science and Engineering Department and the School of Information Technology of IIT Kharagpur. 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.

Niloy Ganguly
Email: niloy@cse.iitkgp.ernet.in

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

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.

Partha Bhowmick

Email: pb@cse.iitkgp.ernet.in

Research areas: *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.

Pallab Dasgupta

Email: pallab@cse.iitkgp.ernet.in

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

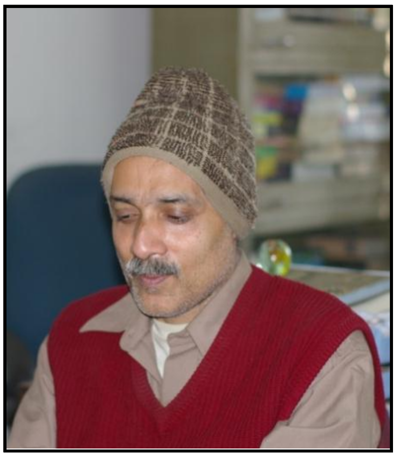
Dr. Pallab Dasgupta did his B.Tech, M.Tech and PhD in Computer Science from the Indian Institute of Technology Kharagpur. He is currently a Professor at the Dept. of Computer Sc. & Engg, I.I.T. Kharagpur. His research interests include Formal Verification, Artificial Intelligence and VLSI. He has over 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 Pratim Chakrabarti

Email: ppchak@cse.iitkgp.ernet.in

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 recepient 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.



Partha Sarathi Dey

Email: psd@cse.iitkgp.ernet.in

Research Interest: *Digital logic design, data structures, computer organization and architecture*

M.Tech.(IIT Kharagpur)
Lecturer, Computer Science & Engineering
P S Dey joined the Institute in 1985

Pabitra Mitra

Email: pabitra@cse.iitkgp.ernet.in

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.

Rajat Subhra Chakraborty
Email: pabitra@cse.iitkgp.ernet.in

Research Interests: *Hardware Security, VLSI Design and Digital Content Protection through Watermarking*

Rajat Subhra Chakraborty is an Assistant Professor in the Computer Science and Engineering Department of IIT Kharagpur. He received his PhD degree in Computer Engineering from Case Western Reserve University (Cleveland, Ohio, USA) in 2010 and a B.E. (Hons.) degree in Electronics and Telecommunication Engineering from Jadavpur University in 2005. From 2005-2006, he worked as a CAD Software Engineer at National Semiconductor in Bangalore, and in Fall 2007, he was a co-op at Advanced Micro Devices(AMD) in Sunnyvale, California. He has received multiple student awards from IEEE and ACM, and an annual award for academic excellence among graduate students from Case Western Reserve University in 2009. Part of his PhD research work has been the subject of a U.S. patent filed by Case Western Reserve University in 2010. His research interest includes hardware security, including design methodology for hardware IP/IC protection, hardware Trojan detection/prevention through design and testing, attacks on hardware implementation of cryptographic algorithms and digital watermarking.

Rajeev Kumar

Email: rkumar@cse.iitkgp.ernet.in

Research Interest: *Programming Languages & Software Engineering, Embedded & Multimedia system, Evolutionary Computing*

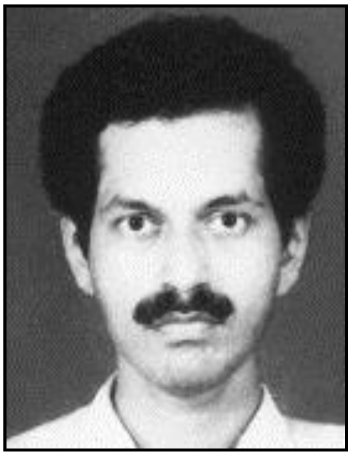
Rajeev Kumar received his Ph.D. from University of Sheffield and M.Tech. from University of Roorkee (now, IIT Roorkee) both in computer science and engineering. Currently, he is a professor of computer science and engineering at IIT Kharagpur. Prior to joining IIT, he was with the Birla Institute of 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.

Rajib Mall

Email: rajib@cse.iitkgp.ernet.in

Research Interest: *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.



Sudebkumar Prasant Pal
Email: spp@cse.iitkgp.ernet.in

Research Interest: *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.

Sudeshna Sarkar

Email: sudeshna@cse.iitkgp.ernet.in

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.

Sujoy Ghose

Email: sujoy@cse.iitkgp.ernet.in

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.

Publications by
Current Research Scholars

2011

1. A.Sarkar, S. Ghose, P.P. Chakrabarti, "Sticky-ERfair: A Task-Processor Affinity Aware Proportional Fair Scheduler," *Real Time Systems Journal*, Springer. Accepted in Jan, 2011.
2. Antara Ain, Debjit Pal, Pallab Dasgupta, Siddhartha Mukhopadhyay et. al, "Chasis: A Platform for Verifying PMU Integration using Auto-generated Behavioral Models", accepted in *ACM TODAES*
3. C. Rebeiro and D. Mukhopadhyay, "Cryptanalysis of CLEFIA using Differential Methods with Cache Trace Patterns", *CT-RSA 2011* (accepted).
4. D. P. Dogra, K. Nandam, A. K. Majumdar, S. Sural, J. Mukhopadhyay, B. Majumdar, S. Mukherjee, A. Singh, "A Tool for Automatic Hammersmith Infant Neurological Examination," *International Journal of E-Health and Medical Communications* Accepted for publication in the 2nd issue of 2011.
5. S. Bandopadhyay, R. Naskar and R. S. Chakraborty, "Reversible Digital Watermarking using Integer Wavelet Transform", *International Conference on Scientific Paradigm Shift In Information Technology & Management*, Kolkata, 2011.
6. S. Das, and P. Mitra, "A Rule-based Approach of Stemming for Inflectional and Derivational Words in Bengali", *IEEE TechSym 2011*, Kharagpur, January 14-16, 2011.
7. S. Das, S. Banerjee, and P. Mitra, "Anwesan: A Search Engine for Bengali Literary Works", *International Conference on Digital Library Management (ICDLM) 2011*, Kolkata, January 11-13, 2011.
8. S. Ghosh, A. Srivastava and N. Ganguly, "Assessing the Effects of a Soft Cut-off in the Twitter Social Network", *IFIP Networking Conference 2011*, Valencia, Spain, May 2011 (accepted).
9. S. Ghosh, G. Korlam, A. Srivastava, and N. Ganguly, "Effects of a Soft Cut-off on Number of Links in the Twitter Online Social Network", *Microsoft TechVista PhD Poster Competition*, Pune, India, (awarded 2nd prize in the competition)January 2011 .
10. S. Ghosh, G. Korlam, and N. Ganguly, "Spammers' Networks within Online Social Networks: A Case-Study on Twitter", *ACM World Wide Web Conference (WWW)*, Hyderabad, India, March 2011 (accepted).
11. S. Ghosh, P.Kane, and N. Ganguly, "Identifying Overlapping Communities in Folksonomies or Tripartite Hypergraphs", *ACM World Wide Web Conference (WWW)*, Hyderabad, India, March 2011 (accepted).
12. S. Roy, B. B. Bhattacharya, P. P. Chakrabarti and K. Chakrabarty, "Layout-Aware Solution Preparation for Biochemical Analysis on a Digital Microfluidic Biochip", in *Proc. IEEE International Conference on VLSI Design*, Chennai, India, January 2-7, 2011.
13. S. S. Ali, R. S. Chakraborty, D. Mukhopadhyay, and S. Bhunia. "MultilevelThreats: an Emerging Attack Model for Cryptographic Hardware." *DATE 2011*(accepted).
14. S. S. Roy, C. Rebeiro, and D. Mukhopadhyay, "Theoretical Modeling of the Itoh-Tsujii Inversion Algorithm for Enhanced Performance on k-LUT based FPGAs", *DATE 2011*(accepted) .
15. S. Saha and R. Kumar. "Improvement of Bounded-Diameter MST Instances with Hybridization of Multi-Objective EA". In *Proc. ICCCS2011, 2011,*

16. Satya Gautam Vadlamudi, Sandip Aine, and Partha Pratim Chakrabarti, "MAWA*—A Memory-Bounded Anytime Heuristic-Search Algorithm", (accepted in) *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics* 2011.

17. Soumyajit Dey, Dipankar Sarkar, Anupam Basu, "A Kleene Algebra of Tagged System Actors", (accepted in) *IEEE Embedded Systems Letters*, (to be published in) Vol. 3, No. 1, 2011.

18. Soumyajit Dey, Praveen Rokkam, Anupam Basu, "Modeling and Analysis of Embedded Multimedia Applications using Colored Petri Nets" (accepted in) *International Journal of Modeling Simulation and Scientific Computing*, World Scientific Publishing, (to be published in) Vol. 2, No. 2, 2011.

19. Sourya Bhattacharyya, Jayanta Mukhopadhyay, Arun Kumar Majumdar, Bandana Majumdar, Arun Kumar Singh and Chanchal Saha, "Automated Burst Detection in Neonatal EEG" in *International Conference on Bio-inspired System and Signal Processing (BIOSIG)*, Rome, Italy on January, 2011

20. Sourya Bhattacharyya, Subhabrata Ghoshal, Arunava Biswas, Jayanta Mukhopadhyay, Arun Kumar Majumdar, Bandana Majumdar, Suchandra Mukherjee and Arun Kumar Singh, "Automatic Sleep Spindle Detection in Raw EEG Signal of Newborn Babies" in *International Conference on Electronics Computer Technology (ICECT)*, Kanyakumari, India, on April, 2011.

21. Subhadip Kundu, Santanu Chattopadhyay, Indranil Sengupta and Rohit Kapur, “Multiple fault diagnosis based on multiple fault simulation using Particle Swarm Optimization”, in *24th IEEE International Conference on VLSI Design*, 2011.

22. Subhankar Mukherjee and Pallab Dasgupta, “Auxiliary State Machines and Auxiliary Functions: Constructs for Extending AMS Assertions”, in I*nternational conference on VLSI design*, 2011.

23. Sumit Das, Anupam Basu, Sudeshna Sarkar, "Prenominal Modifier Ordering in Bengali Text Generation", (accepted in) *International Conference on Intelligent Text Processing and Computational Linguistics (CICLing 2011)*, Tokyo, Japan, February 2011.

2010

1. A. Hazra, P. Ghosh, P. Dasgupta, and P. P. Chakrabarti, “Coverage Management with Inline Assertions and Formal Test Points”, In *23rd International Conference on VLSI Design,* pp. 140-145, January 2010.
2. A. Hazra, P. Ghosh, P. Dasgupta, P. P. Chakrabarti, "Coverage Management with Inline Assertions and Formal Test Points," In of *VLSI Design Conference* 2010.
3. A. Sarkar, A. Shanker, S. Ghose, P.P. Chakrabarti, “A Low Overhead Partition-Oriented ERfair Scheduler for Hard Real-Time Embedded Systems”, *IEEE Embedded Systems Letters.* Oct, 2010.
4. A. Sarkar, P.P. Chakrabarti, S. Ghose, “Partition Oriented Frame-Based Fair Scheduler”, *Journal of Parallel and Distributed Computing (JPDC)*, vol. 70, no. 7, Elsevier, pp. 707-718, 2010.
5. A. Sarkar, R. Nanda, S. Ghose and P. P. Chakrabarti, “Safe-ERfair - A priori Overload Handling in Fair Scheduled Embedded Systems”. Accepted in: *23rd International Conference on VLSI Design* held jointly with *9th International Conference on Embedded Systems Design (VLSID '10),* 3-7th Jan 2010, Bangalore, India.
6. A.Hazra, S. Mitra, P. Dasgupta, A. Pal, D. Bagchi, and K. Guha, “Leveraging UPF-extracted Assertions for Modeling and Verification of Architectural Power Intent”, In the *47th Proceedings of Design Automation Conference (DAC),* pp. 773-776, June 2010.
7. B. Mitra, A. K. Dubey, S. Ghose, N. Ganguly, “Formal Understanding of the Emergence of Superpeer Networks: A Complex Network Approach,” *ICDCN 2010,* Kolkata, January 2010.
8. B. Mitra, A. K. Dubey, S. Ghose, N. Ganguly, “How do Superpeer Networks Emerge?”, *IEEE INFOCOM 2010,* San Diego, USA, March 2010.
9. C. Karfa, D. Sarkar, C Mandal, “Data-flow Driven Equivalence Checking for Verification of Code Motion Techniques” in *IEEE Computer Society Annual Symposium on VLSI (ISVLSI) 2010,* Lixouri Kefalonia, Greece, pp 428-433, 2010.
10. C. Karfa, D. Sarkar, C. Mandal, “Verification of Datapath and Controller Generation Phase in High-Level Synthesis of Digital Circuits”, *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems,* vol. 29, no. 3, pp. 479-492, 2010.
11. C. Karfa, D. Sarkar, C Mandal, P. Kumar, “Verification of Datapath and Controller Generation phase in High-level Synthesis”, In *IEEE Transaction on Computer Aided Design on Ics* 2010.
12. C. Karfa, D. Sarkar, C Mandal. “Verification of Datapath and Controller Generation Phase in High-level Synthesis of Digital Circuits”, in *IEEE Transactions on COMPUTER-AIDED DESIGN of Integrated Circuits,* page 479-492, Vol. 29, No. 3, March, 2010.
13. C. Misra, B. Bhattacharya, and A. Basu, “A new framework to preserve Tagore songs”, *World Digital Libraries,* Vol. 3, No. 1, 2010.
14. C. Misra, B. Bhattacharya, and A. Basu, “A new framework to preserve Tagore songs”, I*nternational Conference on Digital Libraries,* New Delhi, 2010.

15. C. Rebeiro and D. Mukhopadhyay, “Pinpointing Cache Timinga Attacks on AES”, *23rd International Conference on VLSI Design,* 2010.

16. C. Rebeiro, M. Mondal, and D. Mukhopadhyay, "Pinpointing Cache Timing Attacks on AES", in the Proceedings of *23rd International Conference on VLSI Design (VLSID 2010),* pp 306-311, IEEE Computer Society .

17. C. Rebeiro, S. S. Roy, D. S. Reddy, and D. Mukhopadhyay, "Revisiting the Itoh-Tsujii Inversion Algorithm for FPGA Platforms", accepted for publication, *IEEE Transactions on Very Large Scale Integration Systems*, 2010

18. Chandan Karfa, Dipankar Sarkar, Chittaranjan Mandal, “Verification and Synthesis of Digital Circuits: High-level Synthesis and Equivalence Checking”, *LAMBERT Academic Publishing, ISBN* 978-3-8383-9813-6, August 2010.(Book)

19. D. Dash, A. Gupta and A. Bishnu; "Dynamic Maintenance of Support Coverage in Sensor Network"; Parallel Processing Letters, Vol 20, No 2, June 2010

20. D. P. Dogra, K. Nandam, A. K. Majumdar, S. Sural, J. Mukhopadhyay, B. Majumdar, S. Mukherjee, A. Singh, “A User Friendly Implementation for Efficiently Conducting Hammersmith Infant Neurological Examination”, *12th International Conference on E-Health Networking, Application and Services,* Lyon, France, pp. 374-378, 2010.

21. D. P. Dogra, S. Sinha, A. K. Majumdar, S. Sural, J. Mukhopadhyay, B. Majumdar, S. Mukherjee, A. Singh, “Automatic Posture Estimation for Hammersmith Infant Neurological Examination”, *International Symposium on Medical Imaging-Perspectives on Perception and Diagnostics Organized in Conjunction with the Seventh Indian Conference on Computer Vision, Graphics and Image Processing,* IIT Delhi, MA-202, 2010.

22. D. Patra, J. Mukbopadbyay, A. K. Majumdar, B. Majumdar, D. P. Dogra, “Tele-consultation using Clinical Document Architecture in Disease Specific Domains”, 1*2th International Conference on E-Health Networking, Application and Services*, Lyon, France, pp. 187-194, 2010.

23. Debmalya Sinha, Sandipan Mandal, Anupam Basu, "Sahaj Linux", *IEEE Students Technology Symposium (TechSym 2010)*, Kharagpur, India. April 2010.

24. J. Chandra and N. Ganguly, ”On Coverage Bounds of Unstructured Peer-to-Peer Networks”, In *European Conference on Complex Systems,* September, 13-17, 2010, Lisbon, Portugal

25. J. Chandra, S. Shaw, and N. Ganguly, “HPC5: An Efficient Topology Generation Mechanism for Gnutella Networks”, *Computer Networks*, Volume 54, Issue 9, 17 June 2010.

26. Krishna Kumar S., S. Kaundinya, Subhadip Kundu, and Santanu Chattopadhyay, “Customizing Pattern Set for Test Power Reduction via ImprovedX-identification and Reordering” , *16th ACM/IEEE International Symposium on Low Power Electronics and Design (ISLPED)* 2010.

27. Krishna Kumar S., S. Kaundinya, Subhadip Kundu, and Santanu Chattopadhyay, “Particle Swarm Optimization based Vector Reordering for Low Power Testing”, *IEEE ICCCNT 2010,* Karur, Tamilnadu,2010.

28. M. Desarkar, S. Sarkar and P. Mitra, "Aggregating Preference Graphs for Collaborative Rating Prediction", *ACM Conference on Recommender Systems (RecSys 2010),* Barcelona, 2010.

29. P. Ghosh, and P. Dasgupta, "A Formal Method for Detecting Semantic Conflicts in Protocols between Services with Different Ontologies", *International Conference on Web & Semantic Technology (WeST),* Chennai, 2010.

30. P. Ghosh, and P. Dasgupta, "Detecting Ontological Conflicts in Protocols between Semantic Web Services", *International Journal of Web & Semantic Technology (IJWesT),* Volume 1, Number 4, October 2010.

31. P. K. Bhowmick, A. Basu, P. Mitra, A. Prasad, "Sentence Level News Emotion Analysis in Fuzzy Multi-label Classification Framework," *11th International Conference on Intelligent Text Processing and Computational Linguistics (CICLing),* Accepted for publication, 2010.

32. Pinar Öztürk and Rajendra Prasath, "Recognition of higher-order relations among features in textual cases using random indexing", in Proc. of the *18th Int. Conf. on Case-Based Reasoning (ICCBR 2010),* Lecture Notes in Computer Science, Vol. 6176, pp. 272-286, July, 2010.

33. Pinar Öztürk, Rajendra Prasath and Hans Moen, "Distributed Representations to detect Higher Order Term Correlations in Textual Content", Proc. of the *7th Int. Conf. on Rough Sets and Current Trends in Computing (RSCTC 2010),* Lecture Notes in Computer Science, Vol. 6086, pp. 740-750, June 2010.

34. R. Pal, P Mitra, and J. Mukherjee, "Visual saliency and node centrality measures," proc. of *National Conference on Computer Vision, Pattern Recognition, Image Processing and Graphics*, January 2010.

35. R.Rajendra Prasath and Sudeshna Sarkar, "Unsupervised Feature Generation using Knowledge Repositories for Effective Text Categorization (ECAI 2010)", in Proc. of the *19th European Conference on Artificial Intelligence*, Lisbon, Portugal, pp. 1101 - 1102, August 2010.

36. Rajendra Prasath and Pinar Öztürk, "Similarity Assessment through blocking and affordance assignment in Textual CBR", in Proc of the *Reasoning from Experiences on the Web (WebCBR 2010),* pp. 151-160, July 2010.

37. Ratnajit Mukherjee, Soumyajit Dey, Sumit Das, Anupam Basu, "An Iconic and keyboard based Communication Tool for People with Multiple Disabilities", *IEEE Students Technology Symposium (TechSym 2010),* Kharagpur, India. April 2010.

38. Ritwika Ghose, Tirthankar Dasgupta, and Anupam Basu, "Architecture of a Web Browser for Visually Handicapped People", *IEEE Students' Technology Symposium(TechSym),* pp.325 - 329, 2010.

39. S. Bag and G. Harit, "A Medial Axis Based Thinning Strategy and Srtructural Feature Extraction of Character Images", *International Conference on Image Processing (ICIP),*Hong Kong, 2010.

40. S. Bag and G. Harit, "A Medial Axis Based Thinning Strategy for Character Images," *Second National Conference on Computer Vision, Pattern Recognition, Image Processing and Graphics,* Jan 15-17, 2010, Jaipur, India.

41. S. Ghosh, A. Banerjee, N. Sharma, S. Agarwal, A. Mukherjee, and N. Ganguly, "Structure and Evolution of the Indian Railway Network", *Summer Solstice International Conference on Discrete Models of Complex Systems (SOLSTICE),* Nancy, France, June, 2010.

42. S. Ghosh, G. Korlam, and N. Ganguly, “The Effects of Restrictions on Number of Connections in OSNs: A Case-Study on Twitter”, *Workshop on Online Social Networks (WOSN),* Boston, USA, June, 2010.

43. S. K. Dandapat, B. Mitra, N. Ganguly, “Flexible Load Balancing in Wireless Mobile Environment,” *PhD Forum, ICDCN 2010*, Kolkata, India, January 2010.

44. S. Karmakar, D. Mukhopadhyay and D. Roy Chowdhury. “CAvium - Strengthening Trivium using Cellular Automata”. *Automata 2010*, Nancy, France, June 2010.

45. S. Karmakar, D. Mukhopadhyay and D. Roy Chowdhury. “Cube Attack on a Simplified version of Trivium”. *National Workshop of Cryptology 2010*, Coimbatore, India.

46. S. Karmakar, D. Mukhopadhyay and D. Roy Chowdhury. “d-monomial Tests on Nonlinear Cellular Automata for Cryptographic Design”. *ACRI 2010*, Ascoli Piceno, Italy, September 2010.

47. S. Mohanta, G. Vinod, A. K. Ghosh, and R. Mall, "An Approach for Early Prediction of Software Reliability", *SIGSOFT Software Engineering Notes, ACM*, vol. 35(6), pp. 1–9, November 2010.

48. S. Roy, B. B. Bhattacharya and K. Chakrabarty, "Optimization of Dilution and Mixing of Biochemical Samples using Digital Microfluidic Biochips", *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 29(11), November, 2010.

49. S. Roy, D. Mitra, B. B. Bhattacharya and K. Chakrabarty, "Pin-Constrained Designs of Digital Microfluidic Biochips for High-Throughput Bioassays", in Proc. *IEEE International Symposium on Electronic System Design (ISED)*, Bhubaneswar, India, December 20-22, 2010.

50. S. S. Ali and D. Mukhopadhyay. “Acceleration of Differential Fault Analysis ofthe Advanced Encryption Standard Using Single Fault.” *Cryptology ePrint Archive: Report 2010/451,*2010.

51. S. S. Ali, D. Mukhopadhyay, and M. Tunstall. “Differential Fault Analysis of AES using a Single Multiple-Byte Fault.” *Cryptology ePrint Archive: Report 2010/636,*2010.

52. S. Saha and M. Aslam and R. Kumar. “Assessing the Performance of Bi-objective MST for Euclidean and Non-Euclidean Instances.” In Proc. *Third International Conference of Contemporary Computing, 229--240, Publisher: Springer,*2010.

53. Sandipan Mandal, Biswajit Das, Pabitra Mitra, “ SRUTI-II : A Vernacular Speech Recognition System in Bengali and an Application for Visually Impaired Community”, *IEEE TechSym,*April ,2010.

54. Sanjay Chatterji; Sudeshna Sarkar; Anupam Basu, “Implementation of a Bengali Morphological Generator”, in Proceedings of *8th International Conference on Natural Language Processing (ICON-2010*), pp 300—305,Kharagpur, India, December, 2010.

55. Soumya V.B., Monojit Choudhury, Kalika Bali, Tirthankar Dasgupta and Anupam Basu, “Resource Creation for Training and Testing of Transliteration Systems for Indian Languages”, I*nternational Conference on Language Resources and Evaluation (LREC),* pp. 2902-2905, 2010.

56. Soumyajit Dey, Dipankar Sarkar, Anupam Basu, "A Tag Machine based Performance Evaluation Method for Job-Shop Schedules", *IEEE TCAD,* Vol. 29, No. 7, 2010.

57. Subhadip Kundu, Krishna Kumar S., and Santanu Chattopadhyay, “Test Power Reduction with Test-Time Trade-off”, *IEEE International Symposium on Circuits and Systems, ISCAS,*2010.

58. Sumit Das, Anupam Basu, Sudeshna Sarkar, "Discourse Marker Generation and Syntactic Aggregation in Bengali Text Generation", *IEEE Students Technology Symposium (TechSym 2010)*, Kharagpur, India. April 2010.

59. T. Dutta, D. P. Dogra, B. Jana, “Object Extraction Using Novel Region Merging and Multidimensional Features”, *4th Pacific-Rim Symposium on Image and Video Technology*, NTU, Singapore, pp. 356-361, 2010.

60. Tirthankar Dasgupta, Anupam Basu and Pabitra Mitra, “A Framework for the Automatic Generation of Indian Sign Language”, *Journal of Intelligent Systems (JIS)*, 2010, pp. 125-144

61. Tirthankar Dasgupta, Anupam Basu, Animesh Das and Promothesh Mondol, “Design and Evaluation of Bangla Keyboard Layouts”, *IEEE Students’ Technology Symposium (TechSym)*, pp. 248 - 254,2010.

62. Tirthankar Dasgupta, Monojit Choudhury, Kalika Bali, Anupam Basu, "Mental Representation and Access of Polymorphemic Words in Bangla: Evidence from Cross-modal Priming Experiments", International Conference on Natural Language Processing (ICON), pp. 58-67, 2010.

```
g state information... 
llowing extra packages will be installed:
d9utils
sted packages:
anlubis@linux-ubuntu:~$ sudo apt-get install bind9
ding package lists... Done
lding dependency tree
ding state information... Done
 following extra packages will be installed:
bind9utils
ggested packages:
bind9-doc resolvconf
he following NEW packages will be installed:
 bind9 bind9utils
 upgraded, 2 newly installed, 0 to remove and 0 not upgraded.
eed to get 333kB of archives.
After this operation, 1028kB of additional disk space will be used.
Do you want to continue [Y/n]?
```

INDIAN INSTITUTE OF TECHNOLOGY
KHARAGPUR -721302