60

1951-2011

IITKharagpur

Diamond Jubliee 2011-12

RSD

Research Scholars' Day

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.

Like the previous year, this year also we are holding our research scholar day on the 17th March, 2012. This is the third occasion when we, all, will assemble together and get a glimpse of our current research activities from our students. No doubt this year too the day will be observed with equal enthusiasm and zeal by our PhD and MS students, who would take this opportunity to demonstrate their latest research findings and to exchange ideas of research, development, and knowledge among the different research groups and faculties of the department. My best wishes are with them in this endeavor.

Jayanta Mukhopadhyay

(Head)

PhD & MS SCHOLARS

List of Current PhD & MS Scholars

(Arranged According to the Date of Joining)

PhD Scholars

1. Srobona Mitra
2. Priyankar Ghosh
3. Rajiv Ranjan Suman
4. Sanjay Chatterji
5. Dinesh Dash
6. Maunendra Sankar Desarkar
7. Soumen Bag
8. Sk Subidh Ali
9. Rajendra Prasath R
10. Prasenjit Mondal
11. Soumyadip Bandyopadhyay
12. Ishani Chakraborty
13. Bodhisatwa Mazumdar
14. Saptarshi Ghosh
15. Rajib Ranjan Maiti
16. Manjira Sinha
17. Chhabi Rani Panigrahi
18. Soma Saha
19. Sourav Kumar Dandapat
20. Kamalesh Ghosh
21. Chester Rebeiro
22. Satya Gautam Vadlamudi
23. Sudip Roy
24. Rishiraj Saha Roy
25. Sumanta Pyne
26. Bibhas Ghoshal
27. Subhasish Dhal
28. Kunal Banerjee
29. Tirthankar Dasgupta
30. Aritra Hazra
31. Subhadip Kundu
32. Parantapa Bhattacharya
33. Ruchira Naskar
34. Sabyasachi Karati
35. Mahesh Shirole
36. Sudipta Saha
37. Sandip Karmakar
38. Joy Chandra Mukherjee
39. Sudakshina Dutta
40. Tripti Swarnkar
41. Sanjoy Pratihar
42. Tanmoy Chakraborty
43. Anupam Mandal
44. Durga Prasad Sahoo
45. Tanwi Mallick

MS Scholars

1. Biswanath Barik
2. Debjit Pal
3. Anup Kumar Bhattacharya
4. Satrajit Ghosh
5. Praloy Kumar Biswas
6. Sirsendu Mohanta
7. Sourya Bhattacharyya
8. Sandipan Mandal
9. Biswajit Das
10. Ritwika Ghose
11. Prosenjit Dhole
12. Animesh Srivastava
13. Debmalya Sinha
14. Biswanath Saha
15. Chandan Misra
16. Rajdeep Mukherjee
17. Binanda Sengupta
18. Partha De
19. Suman Kalyan Maity
20. Suvadeep Hajra
21. Indrasish Saha
22. Parnab Kumar Chanda
23. Tamal Sen
24. Ayan Palchaudhuri
25. Arnab Dhar
26. Souvik Kolay
27. Debapriya Basu Roy

Research Abstract of PhD & MS Scholars

PhD Scholars

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 March 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 Semi-Formal, Formal and Post-Silicon Verification of Hardware Designs .*

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

Formal Methods for Effective Verification of Local Design Changes

Large digital integrated circuits typically consist of multiple functional units integrated using glue logic. Design teams spend a considerable amount of effort in ensuring the correctness of the main functional units in the design, where current practice includes the use of formal methods. On the other hand, the glue logic, also called tribal logic, which accounts for a large fraction of the integrated circuit (40% or above), is often developed from an intuitive understanding of the architecture, changes significantly from one design to another, is typically not documented properly and typically not formally verified in industrial practice. The size of the glue logic is significant and can easily consist of many components having both sequential and combinational elements.

The glue logic is often modified locally during the design cycle with the expectation of local side-effects. These local design changes in the glue logic can typically be applied for one of the following two reasons:

(a) *Bug Fix:* Bugs appearing in the glue logic are often due to incorrect interpretation of the architecture in specific corner case scenarios which were overlooked during simulation. Such bugs often escape pre-silicon validation and are uncovered during post-silicon testing. When such a bug is detected, the appropriate place in the glue logic is corrected.

(b) *Intent Modification:* The design intent of a component is revised later in the design cycle and accordingly the component is modified locally for correct implementation of the design intent.

However, these local changes may indirectly affect a much larger portion of the glue logic, and it is non-trivial to determine the exact boundary of the cone-of-influence of that change. The task of using formal methods to verify whether the functionality of the modified glue logic in its entirety remains unchanged even after application of these local design changes, is a very hard problem in practice considering the enormity of the glue logic and the nature of the cone of variables influencing the local design change. Model checking or sequential equivalence checking on the entire glue logic as a whole, does not scale. In this work, we propose methodologies for effectively verifying such local design changes at a global level, without attempting to apply formal methods on the entire glue logic. Specifically we solve the following problems:

- **Verification by parts: reusing component invariant checking results:** This work explores the utility of reusing proven component invariants in the backward reachability-based sequential equivalence checking paradigm of formal verification, for verifying that the modified glue logic is equivalent to the previous version. We present a formal method for simplifying the process of proving global invariants on an integrated design (total glue logic) using the reachability information of the component state spaces, obtained from known invariants for the components of the design which remain unchanged. Experimental results on benchmark circuits reveal that deriving the approximate reachability don't cares from the proofs of component invariants helps in reducing both the depth and breadth of the search.
- **Trace Assisted Formal Methods for the Verification of Bug Fixes:** Bug traces reproduced in simulation serve as the basis for patching the RTL code which is essentially a local design change in order to fix the bug. It is important to prove that the patch covers all instances of the bug scenario, failing which, the bug may return with a different valuation of the variables involved in the bug scenario. This work proposes formal methods inspired from software debugging for analyzing the control trace, obtained from the given bug trace, leading to the observed manifestation of the bug and verifying the robustness of the bug fix with respect to that control trace. Our methods provide formal guarantees with respect to the specific bug scenario, which are more scalable by orders of magnitude than model checking the entire design. We believe that the proposed formal methods hold immense promise in analyzing bug fixes in large industrial strength designs.
- **Formal Methods for Ranking Counterexamples Through Assumption Mining:** Verifying local design changes on the total glue logic involves cutting out a cone-of-influence of the change and verifying the property on that cone-of-influence in isolation typically throws up a large number of counterexamples, many of which are spurious because the scenarios they depict are not possible in the entire logic. In this work, we introduce the notion of ranking the counterexamples so that only the most likely counterexamples are presented to the designer. Our ranking is based on assume properties mined from existing simulation traces of the entire logic and also the already proved properties for different modules of the glue logic. We define a metric to compute a belief for each assume property that is mined, and rank counterexamples based on their conflicts with these properties. Experimental results demonstrate an amazing correlation between the real counterexamples (if they exist) and the proposed ranking metric, thereby establishing the proposed method as a very promising verification approach.

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

Priyankar Ghosh

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

Priyankar Ghosh *received a 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.*

Supervisor: 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.

- **Usage of High Level Model in Protocol Verification:** Web service providers typically publish a high level model of the service to describe the behavior of the web service. Typically these models are written in English language and may have some graphical representation as well. In this work our goal is to formalize the model, and generate a set of test cases using the model in order to verify the correctness of the implementation. These test cases will include positive as well as negative test cases. This published model is also used during the integration with the client. During the integration with the server, the client side writes test cases to check the protocol compliance. Diagnosing the reason of the mismatch in message exchange between the client and the server plays a crucial role in debugging the client applications. Our goal is to develop and formalize a debug mode based testing methodology which will assist the client to detect reason of interaction failure during the integration with server and provide useful information regarding the test cases developed by the client.

- **Modeling Semantic Information Exchange and Detecting Conflicts:** Interaction between the client and the server may also fail due to the difference in the interpretation of the exchanged data. The semantics of data play a major role in semantic web services which goes beyond simple type checking. Therefore the protocol that defines the interaction has to be verified in order to check the presence of semantic conflict. It is possible that the knowledge base of the server has some conflict with the knowledge base of the client, but the protocol does not sensitize the conflict.

- **Planning Based Approach to Compose Web Services:** It is quite common to use multiple web services in order to achieve a goal. Since the overall goal may be achieved only when these web services are invoked in a particular order, it is often needed to undo the effect of one service. For example hotel reservation may be canceled 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

Email: rrsuman2001@gmail.com
Joined the department in: July 2007

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

Supervisor: Prof. Rajib Mall

Construction of State Models of Software Components

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

In component-based software development paradigm, a large software is built by assembling pre-built and independently developed “plug and play” type of executable units, called software components. Only a brief description of the functionality of a component is provided by the component vendor. 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.

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

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

State models (FSMs, statecharts, etc) of objects in object-oriented systems is a behavioral model that depicts the different states that the object may assume and transitions among the states that may occur in response to the stimuli received from its environment. Externally, the state model of a component is visualized in terms of the state-based behavior of the component as a whole rather than in terms of the individual objects that the component may be composed of.

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

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

Sanjay Chatterji

sanjaychatter@gmail.com
Joined the department in: January 2008

***Sanjay Chatterji** received a B. Tech. degree in Computer Science and Engineering from the Haldia Institute of Technology in 2003, and Master of Engineering degree in Computer Science and Technology from Bengal Engineering and Science University (Formerly 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 Machine Translation and other areas of Computational Natural Language Processing.*

Supervisors: Prof. Sudeshna Sarkar and Prof. Anupam Basu

Bengali Hindi Machine Translation

Machine translation is a process by which a text from one language (source) is translated to a text in another language (target). There are two major paradigms of machine translation: Statistical Machine Translation and Rule Based Machine Translation. Statistical machine translation builds a model based on the bilingual parallel sentences. On the other hand, rule based system is built based on dictionary entries and analysis of source language text. Different levels of syntactic and semantic analysis may be used in the rule based approach.

The objective of our work is to develop Bengali Hindi machine translation system. 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. Even though Bengali and Hindi are close language pairs and have a great degree of syntactic similarity, we observe that a lexical level based transfer of Bengali text does not always output correct and fluent Hindi text. We have prepared a baseline transfer based translation system using morphological analysis, part of speech tagging and chunking of the source sentence, grammatical and lexical transfer of source sentence to target sentence and generation of words in the target language.

This transfer based approach has given us a basic translation system that produces basic translation. But, many errors and fluency issues remain. The solutions to some of the problems require deep syntactic analysis of the source sentence, and in some cases semantic information becomes essential. However a transfer based system has many shortcomings, some of which are better addressed by a statistical translation system. We also developed a basic statistical machine translation system, though we are limited by the resources. This system is able to handle certain constructs better but does not work for all structures. One of the objectives of our work is to identify constructs that can be best handles by one of the methods and to develop a good quality hybrid machine translation system between Bengali and Hindi.

We have identified some issues in transfer based machine translation from Bengali to Hindi that are being solved in our work. For example, the copula verbs in present tense and sometimes in past tense are dropped in Bengali but not in Hindi. So, we have to find the suitable copula to be used and the place where it will be placed. In Bengali to Hindi translation inter chunk reordering is rare, but inside a chunk some words swap their places when translated. Another issue is the translation of the particles. Some particles are dropped in Hindi and some are translated with a change in the whole sentence structure. There are some agreement related issues that require long distance relations. Syntactic processing will be essential to solve this. In some cases the appropriate translation of a word or a phrase depends on its context. In cases where the context is local, statistical systems can find a good solution to the problem, but where the context is long distance, syntactic parsing may be essential. We have also improved the lexical selection of Bengali Hindi transfer based system with a statistical model.

Nowadays, statistical approach is widely used for machine translation. With limited amount of resources we have developed some baseline Bengali Hindi statistical translation systems using some open source systems. Due to the lack of coverage some words are not translated to target language. We have tried to translate those 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. This output is further processed by applying some affix based postprocessing rules. Some of the words are wrongly translated. Therefore, the translations are revisited by language dependent hard constraints.

Further, the final output of both the hybrid systems can be modified using rules and corpus. Creation of the rules and parallel corpus will take large amount of time. The monolingual corpus of target language is crawled from the web and used to modify the system outputs. First, we have used modified beam search algorithm and finally a hybrid approach of rules and modified CKY algorithm. We have also addressed the problems of Treebank, Parsing, Anaphora resolution, Morphological analysis and synthesis, etc.

Dinesh Dash

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

Dinesh Dash *received a 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 algorithms in Wireless Sensor Network.*

Supervisor: Prof. Arobinda Gupta and Prof. Arijit Bishnu

Geometric algorithms for coverage problem in wireless sensor network

One of the fundamental issues in the deployment of a Wireless Sensor Network is the ability of the network to sense environmental parameters such as temperature, pressure etc. 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. It is important to measure how well the region is sensed by the deployed sensors. These measures are called the quality of coverage. There are different types of coverage measures depending on the application of the sensor network. *Area coverage* [1] ensures that every point of the selected region is under the sensing range of some specific number of sensors. Instead of covering the whole region if we are interested only to a set of points then there is another measure called *target coverage* [2]. *Target coverage* ensures that all the target points are in the sensing range of at least a given number of sensors. In *barrier coverage* [3] there is a given bounded region and it ensures all the crossing paths through the boundary detected by the sensor network. My research works are focused on three aspects. First, measure the quality of coverage for a given deployment. Second, designing deployment plans to ensure a given quality of coverage. At last, we propose different coverage maintenance schemes to recover the coverage loss due to node failure for power drainage or environmental calamities.

Object tracking is one of the fundamental applications in sensor network. Here the sensors detect a moving object and return its position. In my first work, we assume that the moving objects are walking in straight line path. A line segment is said to be *k- covered* if it intersects at least k sensors' sensing regions and is *k-uncovered* if it intersects at most k-1 sensors' sensing regions. 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. A straight line segment is said to be k-covered if it intersects at least k sensors' sensing regions. In this work, we have determined the *longest k-uncovered* segment and *smallest k-covered* segment within the region.

In my second work a deployment scheme is proposed to ensure *k-line coverage* for the set of line segments. The deployment scheme is such that it uses minimum number of 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 and for fixed length segment of arbitrary orientation.

Sensors are subject to failure and therefore the quality of coverage degrades as time passes. In my final work, we are maintaining the quality of coverage after node failure by moving local redundant sensors. Here we propose coverage maintenance schemes for support coverage and barrier coverage after sensors failure.

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.

Maunendra Sankar Desarkar

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

***Maunendra Sankar Desarkar** received a 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. He received the Microsoft Research India PhD fellowship from Microsoft Research India in 2009. His research interests are in the areas of Data Mining and Information Retrieval.*

Supervisor: Prof. Sudeshna Sarkar

Algorithms for Personalized Recommender Systems

Recommender systems suggest to the users different items that the users might be interested in. The items can be of various types such as movies, music, books etc. Often different users have different preferences for these product types. While recommending items to a user, it becomes important to understand the choice of the user and suggest items according to his/her taste. We focus 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 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.

Neighborhood based collaborative filtering is a widely used framework for rating prediction in recommender systems. Based on the assumption that users with similar tastes would rate items similarly, this framework first finds a group of users having similar interests. Ratings given by the users from that group are used to predict unknown ratings. User-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.

We developed a *preference relation based collaborative filtering algorithm* for the Rating Prediction problem. Collaborative systems looking 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 the 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 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. Experimental results show that our method outperforms other methods in the sparse regions.

We also look at the problem of *item recommendation using preference relations in the matrix factorization framework*. From the absolute ratings provided by the users, we induce preference relations and feed those to the proposed algorithm. The algorithm first models each user and each item as a point in a low dimensional feature space. Each dimension can be viewed as a hidden category mined from the data. The low dimensional feature representation of an item suggests the item’s belongingness to those latent categories. Similarly, the feature representation of a user suggests the user’s affinities to those categories. Computation of the user and item features is performed offline. Once the feature representations are available, the system may predict the items’ utilities to different users. This information is then used to generate the recommendations. This recommendation generation part is performed online when the user accesses the system or explicitly asks the system to recommend items for him/her. Experiments performed on a benchmark dataset show that the proposed method is able to achieve better recommendation accuracy compared to the alternative algorithms.

We also study recommender algorithms that consider *temporal information for item recommendation*. We look at ways of incorporating purchase time information in the standard user based collaborative filtering algorithms. Users’ interests may shift over time. Recommender systems should therefore rely on recent purchases of the users. Items also have their own dynamics. Most of the items in a recommender system are widely popular just after their releases but do not sell that well afterwards. The proposed algorithms use the time-of-purchase information for calculating user similarities. The time information is also used while combining the *purchase behaviors* of the *experts* and generating the final recommendation. Experimental comparisons performed on several benchmark datasets indicate that the recommendation performance can be improved by considering the recent purchases of the users and the experts.

Soumen Bag

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

***Soumen Bag** received the B.E. and M.Tech. degree in Computer Science and Engineering from NIT Durgapur, India, in 2003 and 2008 respectively. From January 2004 to June 2006, he worked as a lecturer in the department of Computer Science and Engineering in BCET Durgapur, India. Since July 2008, he has been a Research Scholar in the department of Computer Science and Engineering in IIT Kharagpur, India. His research interests are in the areas of OCR for Indian Scripts, Document Image Analysis, Image Processing, and Pattern Recognition.*

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

Devising Techniques and Features for High Performance Character Recognition for Bangla Script

Optical character recognition (OCR) has been an active subject of research for several decades. Many OCR systems are available in the market; however their performance degrades significantly with different fonts, orientation, quality, and presence of compound (also known as “conjunct”) characters or composites in the script. Our goal is to improve the performance of existing OCR systems for Bangla which is the second most popular language in India and fifth most popular language in the world. It is also used as a script to write different other Indian languages. So Bangla has a big importance both as a script and as a language. Our research focuses towards developing new features by suitably analyzing the structural shape of characters. Typically, OCR systems for Indian scripts involve elaborate preprocessing steps involving binarization, skew correction, text line, word, and character segmentation, skeletonization, etc. To start with, we propose an adaptive-cum-interpolative binarization method in a multi-scale framework to handle degraded documents, a character segmentation method based on vertex characterization of outer isothetic polygonal covers, and a medial-axis based thinning strategy which is tailored for thinning printed and handwritten characters in Indian scripts.

We present novel topological features based on the structural shape of a character for performing Bangla basic character recognition. By topology of a character, we mean the structural features of the strokes and their spatial relations. Our objective is to formulate features pertaining to the topology of the character. We detect the convex-shaped segments formed by the various strokes. The convex segments are then represented with shape primitives from a repertoire. We formulate feature templates for Bangla characters. A given character is assigned the label of the best matching feature template. We have tested our method on a benchmark datasets of printed and handwritten Bangla basic character images. Our results demonstrate the efficacy of our approach comparing with other Bangla OCR methods.

Finally, we focus on the challenges to recognize compound characters in Bangla script. There is a large number of (near about 250) compound characters in Bangla. Many of them are very complex in shape when compared with compound characters in other Indian languages. The proper recognition of such complex-shaped compound characters is a difficult problem. To make the problem simple, we try to break down the compound characters into shape components (i.e., shape components to be recognized). The novelty in this approach lies in the formulation of character decomposition rules which consider different grouping principles for grouping stroke segments to form shape components. The identity of the compound character is done by taking note of the identity of the shape components using topological features and string matching technique. Our technique is applicable to printed and handwritten compound characters. The proposed method performs well for some complex-shaped compound characters which were confusing to existing methods.

Sk Subidh Ali

Email: subidh@cse.iitkgp.ernet.inwherever.com
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: Dr. 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, Shamir *et.al.* 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^{32} . 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.

There is another kind of fault attack on AES crypto-system where the fault is induced at the key scheduling algorithm. Until recently it was assumed that the DFA on AES key schedule is much more difficult than the normal DFA on AES, where the fault is induced at the states. The most optimum attack on AES-128 key schedule required around two fault induction and a brute-force search of 32 bits. We proposed new attacks on all the three versions of AES which show that the AES key schedule is as vulnerable as the AES state. Our attack on AES-128 key schedule takes only one faulty ciphertext and a brute-force search of 8 bits to reveal the secret key. The attacks on AES-192 and AES-256 take two and three faulty ciphertexts respectively in order to reveal the secret key. We also develop reduction techniques to prove that the proposed attacks are the most optimized attacks possible on AES crypto-system based on fault induction.

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 countermeasures have large overhead, this motivates us to study low-cost and low-overhead countermeasures against fault attacks on AES.

Prasenjit Mondal

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

Prasenjit Mondal *received an 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 November 2011, he worked in Telemedicine Project, IIT Kharagpur, as a Junior Project Officer. From December 2011 till now, he is working in Document Image Analysis Project. 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.*

Supervisor: Prof. Jayanta Mukhopadhyay and Prof. Shamik Sural

Segmentation of Human Brain USG and MR Images

With the advancement of technology, healthcare services have been improved through various image processing techniques. Ultrasound imaging or ultrasonography (USG) is widely used as a non-invasive diagnostic tool to produce images of the internal organs of the body. In the last few decades, research in ultrasound imaging has seen a noticeable progress. For example, automatic object contour detection, object segmentation, feature extraction and noise reduction from ultrasound images are commonly used in medical image processing. Ultrasound imaging is widely used because it is safe, low cost and easily available compared to other imaging techniques such as CT scan, MRI, etc. Magnetic resonance (MR) imaging is a powerful medical imaging technique commonly used to visualize detailed information about soft tissues at a high resolution as compared with other medical imaging techniques such as computed tomography (CT) or X-rays. Image segmentation refers to the partitioning of an image into different regions in a way so that the set of points belonging to the same region share certain visual characteristic.

Brain USG is mainly done for neonates. The quality of brain ultrasound images is quite low. Also, speckle noise present in the images makes it difficult to recognize boundaries of different objects. It needs expertise where one can recognize objects from the brain ultrasound images. Due to high complexity of the anatomical structure of human brain, segmentation of USG and MR images is a challenging problem. Appropriate segmentation of brain components can help researchers and physicians in understanding and analyzing the structure and function of the brain. This has motivated us to develop algorithms to automatically segment different anatomical structures from USG and MR images.

To segment ventricle system (elliptical shape in nature) from USG, a computational model has been developed that automatically fits ellipses to the ventricle system. The boundary of the ellipse is used as the initial boundary of the ventricle system which is further fine tuned by means of active contour model (snakes). To segment MR images, an efficient dynamic programming based technique has been developed that uses active contour model in the last stage of segmentation. The algorithm can segment brain MR images in axial, sagittal and coronal views. It has been observed that the proposed technique can efficiently segment USG and MR images with high accuracy.

Reference

[1] J. M. Rennie, C. F. Hagmann and N. J. Robertson, "Neonatal Cerebral Investigation", Cambridge University Press, 2008.

Soumyadip Bandyopadhyay
Email: soumyadip@cse.iitkgp.ernet.in
joined the Department in: January 2009

Soumyadip Bandyopadhyay *received a B.Tech degree in Computer Science & 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 and Embedded Systems.*

Supervisors: Prof. Chittaranjan Mandal and Prof. Dipankar Sarkar

Behavioural Verification of Embedded Systems using PRES+ model

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 a PRES + model (Petri net based Representation for Embedded Systems)[3] as an extension of classical Petri net model that captures computation, concurrency and timing behaviour of embedded systems; it allows systems to be represented in different levels of abstraction and improves expressiveness by allowing the token to carry information. This modeling formalism has a well defined semantics so that it supports a precise representation of system.

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 optimizing 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. The input beaviour undergoes several transformations steps before being mapped to an architecture. Our objective is to verify those transformation steps.

Specifically, we address two issues —automated checking of functional equivalence of the transformed optimized behavioural specification to the original one, also referred to in the literature as transformation validation, and comparison of the performance of the timing behaviours of the design before and after the optimizations are applied. While the sequential behaviour can be captured by FSMDs, the parallel behaviour can be easily captured using PRES+. An equivalence checker for FSMD models already exists [4].

Hence, we have formulated an algorithm to translate a PRES + model into an FSMD model and use existing FSMD equivalence checker. It is to be noted that timing constraints are inconsequential for demonstrating data transformation equivalence between the behaviours which allows us to perform equivalence checking using FSMDs. However, translation of a PRES+ model into the corresponding FSMD model encounters state explosion because the method essentially involves parallel composition of the concurrent transitions in PRES+. Moreover, the state explosion problem is further aggravated due to various possible interleaving of the concurrent transitions, which may come into play when timing analysis is addressed. Therefore, we have formulated a direct equivalence checking between two PRES+ models. As a part of the future work, a comparative study of the two equivalence checking methods, one using FSMDs and the other using PRES + models, will be evolved. Specifically, we intend to address code motion validation for this comparative study.

Next we aim at the following enhancements of the PRES+ equivalence checker: (i) incorporation of arrays, (ii) capabilities for loop transformation validation and (iii) synchronization transformation. Another goal is to evolve timing analysis mechanism. Some optimizations target timing performance improvement and some resource utilization. Analysis mechanisms will be evolved for both situations. In the former case, the analysis should demonstrate how much the improvement is. In the latter case, it verifies that the deadlines are not compromised.

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, pages 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, pages 132–139, Washington, DC, USA, 1998. IEEE Computer Society.

[3] L. A. Cottez, 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, pages 149–155, Washington, DC, USA, 2000. IEEE Computer Society.

[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, 27(3):556–569, 2008.

Bodhisatwa Mazumdar

Email: bodhisatwa@cse.iitkgp.ernet.in
Joined the department in: July 2009

***Bodhisatwa Mazumdar** received a 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

Design for Security of Block Cipher S-Boxes to Resist DPA Attacks

As communication networks are spreading by leaps and bounds, the need for secured communication and hence fast but secured cryptographic systems is growing bigger. This necessitates the call for “Design for Security” which entails design, implementation and security of cryptographic hardware and embedded systems. Block cipher cryptosystems embedded in cryptographic devices are susceptible to attacks which show that security cannot be an afterthought. Like as performance, testability are important issues which the designer takes care in the design cycle, security also has to be taken into consideration early in the design cycle. As a motivating example, differential power analysis (DPA) attacks and vulnerability modeling of cryptographic devices have shaken the strength of cryptographic implementations and have baffled the designers for the past fifteen years that a very strong mathematical algorithm can be compromised using these powerful techniques. The Advanced Encryption Standard (AES) was found to be compromised, despite being a mathematically strong cipher. Literature shows that S-boxes, the nonlinear components in the cipher are responsible for its vulnerability towards DPA.

The asymmetry in the power consumption of transitions of bit values from 0 to 1, and 1 to 0 in the CMOS library are the root cause for such attacks. Countermeasures like gate-level masking and several algorithmic countermeasures have been discovered but they were found to be costly in terms of hardware footprint and power consumption. Also some of these countermeasures are prone to higher-order differential side channel attacks. Most importantly, they make AES like block cipher algorithms so poor in performance that they rob the algorithm from its mathematical elegance and efficient performance, some of the prime reasons why Rijndael algorithm emerged as the AES.

With this motivation, the current research work investigates whether Boolean functions involved in the AES S-Box can be designed with security as an objective from the beginning. In the first step, we have proposed a RAIN S-Box whose construction and design makes it more DPA resistant than the AES Rijndael S-Box while having similar classical cryptographic properties like SAC, nonlinearity, balancedness, propagation characteristics (PC) and correlation immunity (CI). Also the parameter, transparency order which quantifies DPA resistance of the S-Box is found to be smaller than the Rijndael inverse S-Box and has been practically shown to require more number of power traces to attack the cipher. This confirms that the nature of Boolean functions have strong role not only on the mathematical robustness but also on the resistance to attacks which exploit side-channel information leakage like power. At present, we attempt to synthesize a class of balanced Boolean functions which have both above-mentioned important cryptographic properties and higher resistance DPA attacks defined in terms of parameters like transparency order and SNR (DPA). This involves involving proposing heuristic searching algorithms to find DPA resistive Boolean functions in the class of Rotation Symmetric Boolean Functions which work towards optimising a cost function in terms of Walsh spectra and autocorrelation functions of the coordinate functions of an S-Box.

Saptarshi Ghosh

Email: saptarshi@cse.iitkgp.ernet.in, saptarshi.ghosh@gmail.com
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 Networks, specifically in Online Social Networks and Transportation Networks..*

Supervisor: Prof. Niloy Ganguly

The two faces of the Twitter Social Network
Spammers / Link-farmers and Topical Authorities

The Twitter Online Social Network (OSN) has emerged as one of the most popular social networking sites on the Web. Twitter is presently used not only for communicating with friends and others sharing common interests from all over the world, but also as a platform for discovering real-time information on the Web, such as current events, news stories, and people's opinion about them. Recent estimates suggest that 200 million active Twitter users post 150 million tweets (messages) containing more than 23 million URLs (links to web pages) daily.

As the information shared over Twitter grows rapidly, Twitter *search* is increasingly being used to find interesting trending topics and recent news (as in the Web) [Teevan]. As the methods used to search Twitter become similar to Web search -- such as ranking the users by algorithms such as Pagerank [Talbot] -- it is not surprising that Twitter has started to attract the attention of spammers. Similar to the Web, where some websites exchange reciprocal links with other sites to improve their ranking by search engines, spammers try to infiltrate the Twitter network by link-farming -- they follow other users and try to get others to follow them.

We investigated the vulnerability of the Twitter social network to link farming. Specifically, we attempted to understand the users who establish links to spammers and the potential reasons for their behavior. To this end, we gathered data of 41,352 spammer accounts and conducted a detailed analysis of the users who connect to them. Our analysis revealed surprising social dynamics that drive link farming in Twitter. We found that a majority of links farmed by the spammers come from a relatively small number (100,000 or so) of legitimate, popular, and highly active Twitter users (which is very different from the Web, where popular web-pages would rarely point to spam pages). These users engaging in link farming are *not* spammers themselves, rather they are 'social capitalists', whose goal is to amass social capital and promote their legitimate content on Twitter. Examples of social capitalists range from popular bloggers of social media and Internet technologies to celebrities like Britney Spears and from politicians like Barack Obama to businesses like JetBlue airways. We found that social capitalists tend to connect to anyone who connects to them to increase their social capital. Unfortunately, spammers also exploit this behavior of capitalists to farm links in the Twitter network and promote spam content. We also explored mechanisms to deter link farming in Twitter. We proposed a Pagerank-like ranking system, called Collusionrank, that penalizes users (by lowering their influence scores) for connecting to spammers, thereby disincentivizing users from colluding with unknown people, who might potentially be spammers.

A paper based on the above study (which was conducted in collaboration with researchers at the Max Planck Institute of Software Systems, Germany) has been accepted at the ACM World Wide Web Conference (WWW) 2012 (a preliminary version of the study was also accepted as a poster in WWW 2011).

As the above study showed traditional graph metrics such as number of followers and Pagerank can easily be manipulated in the Twitter network, hence there is a need for better methods to identify authoritative users. In particular, as searching for information becomes more and more common in Twitter, algorithms for identifying topical experts (i.e. users who are authorities on specific topics) need to be developed. We are presently working on this problem of identifying topical experts in Twitter. Prior approaches to identify topical experts in Twitter [Pal] rely either on the information provided by the user herself (e.g. a short bio written by the user, or the tweets posted by the user), or on analyzing the network characteristics of users. We are using an entirely different approach to identify topical experts in Twitter by utilizing crowd sourced topical annotation of experts. For this, we are utilizing the Lists feature in Twitter, using which any user can group Twitter accounts that tweet on a topic that is of interest to her, and follow their collective tweets. We have observed that many users carefully create Lists to include other Twitter users who they consider as experts on a given topic. Furthermore, they generate meta-data, such as List names and descriptions that provide valuable semantic cues to the topical expertise of the users included in the List, which we are utilizing to identify experts in specific topics.

References

1. [Teevan] J. Teevan, D. Ramage, M. Morris, "#TwitterSearch: a comparison of microblog search and web search", ACM WSDM, 2011.
2. [Talbot] D. Talbot, "How Google Ranks Tweets", http://www.technologyreview.in/web/24353/, January 2010.
3. [Pal] A. Pal, S. Counts, "Identifying topical authorities in microblogs", ACM WSDM, 2011.

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

***Rajib Ranjan Maiti** received a B.Sc. (H) degree in Computer Science from Vidyasagar Univeristy, Paschim Medinipur in 2001, an M.C.A degree from Biju Patnaik University of Technology, Orissa in 2004, and an M.Tech. degree in Computer Science and Engineering from Indian Institute of Technology, Kharagpur in 2008. From June 2008 till June 2009, he worked in Magma Design Automation (India) Pvt. Ltd. as an Associate Member of Technical Staff. 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 Peer-to-Peer Networks and Time-Varying Networks.*

Supervisor: Prof. Niloy Ganguly & Prof. Arobinda Gupta

Efficient Routing and Broadcasting in Delay Tolerant Networks

My primary research interest is in designing efficient routing and broadcasting techniques for delay tolerant networks. Delay Tolerant Networks (DTNs) are characterized by the unavailability of end-to-end path between source and destination most of the time due to highly dynamic neighbors and link failures for an unpredictable duration. The communication in DTN is opportunistic since the message is passed only when the devices are in each others' transmission range. I am interested in DTNs where devices are carried and/or controlled by humans. Therefore, in this case, the meeting frequency, in turn the performance of DTN protocols, is primarily controlled by the movement patterns of the humans.

Therefore, initially the performance of the protocols of both routing and broadcasting in DTNs with existing human mobility models has been investigated. Preliminary investigation of human mobility patterns shows that humans do form clusters while moving and the clusters are scattered in the considered geographic location. This motivates me to create some long range contacts which may accelerate a message spreading process by sending a message to a further distance. Directional antenna (DA) is a tool that can help to create such contact. Therefore, the goal is to analyze the impact of using DAs for message transfer in DTNs. Message transfer among agents is carried out using epidemic dynamics. The results show that the use of even a small percentage of DAs along with traditional omnidirectional antenna (OA), placed randomly, can improve the performance of both routing and broadcasting significantly. The impact of different parameters of both antenna and the mobility model used has been investigated. Simulation results show that the performance may be improved irrespective of most of the parameters. Further to this direction, I plan to investigate the impact of placing DAs depending on the mobility behaviors of the agents, subsequently, propose an improved DTN message spreading protocol using DAs.

The existing human mobility models, however, do not produce a generalized practical mobility scenario that is expected to satisfy a certain set of properties reported from various real world trace analysis. Therefore, I have focused on designing a generalized model for human mobility patterns. Preliminary investigation on human mobility patterns reveals that the patterns have various important properties, reported from various real world trace analysis. However, none of the existing mobility models satisfies all the reported properties together. In this part of the work, I have summarized the properties and tried to design a good mobility model that can satisfy these properties. The verification of the proposed model is in progress and the initial results show that it can reproduce human mobility patterns very well satisfying all the summarized properties. Further, I have planned to investigate the impact of each of the properties on the performance of DTN protocols which can further establish the importance of each of the properties in protocol design.

An interesting property of human mobility patterns shows that humans have repetitive patterns of contacts with neighbors and colleagues. This observation indicates that the message spreading can be achieved in an efficient way instead of commonly used probabilistic methods. Therefore, I am also interested to apply the techniques of temporal network theory to design efficient DTN protocols under practical mobility patterns. Subsequently, a theoretical framework for analyzing the effect of temporal properties of human mobility patterns on the protocols of DTN is being developed.

References

[1]. F. Peruani, A. Maiti, S. Sadhu, H. Chat, R. Roy Choudhury and N. Ganguly, “Modeling Broadcasting using Omnidirectional and Directional Antenna in Delay Tolerant Networks as an Epidemic Dynamic”, IEEE JSAC, Vol. 28, 2010.
[2]. K. Lee, S. Hong, S. J. Kim, I. Rhee and S. Chong, “SLAW: A Mobility Model for Human Walks”, INFOCOM, Rio de Janeiro, Brazil, 2009.

Manjira Sinha

Email: manjira87@gmail.com, manjira@cse.iitkgp.ernet.in
Joined the department in: July 2009

***Manjira Sinha** received a 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 Natural Language Processing, Cognitive Science and Text Readability and Enhancement.*

Supervisor: Prof. Anupam Basu

Text Readability and Enhancement

Often, after going through a piece of text, the first criteria by which we judge is the 'readability' of the text. Though we cannot always concretely parameterize, generally it refers to the fact that how well we have grasp the content. 'Readability' is the ease with which a text can be read and understood. It plays a significant role in the design of texts which match to the target populations' skill. Readability has been measured using a number of factors from both the reader's and the text sides. Readability depends on the reader's physical and cognitive abilities as well as social and economic background. For text, readability is generally defined by four top level features: coherence, style, format and organization. Apart from the reader and the text, readability also gets affected by the communicating language. Every language has some unique properties and any effective metric of readability has to take these into account.

English has a long history of readability research starting from 1880. From then, it has come a long way from early subjective evaluation techniques to statistical measures and empirical formulas, from addressing the child population to investigate the reading choices and availabilities for adults. In the beginning, metrics were based on 'vocabulary frequency list', then afterwards, formulas like Flesch index, Fog index etc. incorporated the structural features of a text, and now a days we have measures which takes into account the higher level cognitive features like text cohesion and coherence, e.g. coh-metrix.

In case of Indian languages, especially Bangla, such extensive research work is still unavailable. There are applications of the known readability formulas of English to measure the readability of Bangla texts. The problem in this approach is that Bangla as a language has some distinguishing properties than English, therefore, the formulas applicable to English do not yield the correct results when implanted unchanged in Bangla. Another important aspect of readability, as mentioned above is to customize texts for different reader groups. In the context of a country like India, this is the need of the hour in every level of formal or informal education. If we consider the case of textbooks at the school level, we will see that a majority of both the students and teachers find them extremely hard to comprehend and retain. In addition to this, for a language like Bangla, geographical variations of the language-usage come into play.

To address these and to provide a effective framework for designing textbook contents in Bangla, measures have to be taken at different levels of hierarchy taking into account the backgrounds of both teachers and students; the levels are defined as: 1. The bottom layer will deal with the lexical choices, i.e. which word to use to describe a concept, 2. This layer will analyse the relative difficulties of the different sentence structures, 3. The purpose of this layer will be to study the organization of the discourse and measures of its local and global coherence, 4. At the top level , the balance between diagrams and texts will be considered. As can be seen the bottom two levels will take into account the language specific features and the top two levels will deal with the cognitive and psycholinguistics sides. In this way, it will be possible to have a complete approach towards enhancing the acceptability of textbooks.

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, Odisha 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), Odisha. 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

A Hybrid Regression Test Selection Technique for Object-Oriented Programs

Regression testing is an expensive maintenance activity and is carried out after each modification to a software [1, 2]. Regression Test Selection (RTS) is carried out to ensure that changes do not adversely affect unmodified portions of the software. It often accounts for almost half of the software maintenance costs [1]. RTS techniques help to reduce the time and effort required to carry out regression testing.

RTS techniques based on analysis of both source code and model have been proposed in the literature for object-oriented software. Many RTS techniques first construct either the control flow or the dependency representation of programs based on code analysis and then select test cases. These techniques compare the original and modified versions of the program model and select test cases that execute the affected model elements. However, these techniques ignore the fact that state behavior of objects may get affected due to code changes and the affected state behavior needs to be analyzed to select appropriate test cases. In case of model-based RTS techniques, regression test cases are selected by comparing the original model with the model of the modified program. A problem with this approach is that models being abstraction after all, are often insensitive to minor code changes.

It has been argued that state-dependent behavior of objects is a prominent issue that needs to be addressed during RTS of object-oriented software. A code modification may affect the state behavior of an object. For example, a code modification can lead to a sneak transition path to a state. During regression test selection, the state behavior of an object can not easily be analyzed from code analysis. On the other hand, the state model of an object is usually constructed during the design phase of software development life cycle. Existing code-based RTS techniques for object-oriented programs by and large ignore the state behavior of objects. Unless the state behavior of objects are taken into consideration, several fault-revealing test cases may be omitted during RTS. RTS techniques only based on UML state machine models are limited in the literature and are not precise when compared to code analysis. In this context, we propose an RTS technique that considers control and data dependence information and dependencies arising due to object-relations as well as the state behavior of objects as documented in state models during RTS of object-oriented programs.

In this context, we propose a hybrid regression test selection technique for object-oriented programs. Our proposed technique is based on analysis of both the source code of an object-oriented program as well as the UML state machine models of the affected classes. We first construct a dependency graph model of the original program from the source code. To find the model elements affected due to a program change, we construct a forward slice of the constructed graph model, where each changed model element is used as a slicing criterion. Subsequently, we determine the affected methods from an analysis of the state machine models based on the changed statements. The test cases that exercise the affected model elements in the program model as well as the transitions caused by the affected methods in state machine models are selected for regression testing.

References

[1] G. Kapfhammer. *The Computer Science Handbook*, chapter Software testing. CRC Press, Boca Raton, FL, 2004.

[2] H. Leung and L. White. *Insights into regression testing*. In Proceedings of the Conference on Software Maintenance, pages: 60-69, 1989.

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 was attached with Maharaja Manindra Chandra College, University of Calcutta, Kolkata, as a Guest Lecturer in Department 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 and Prof. Arobinda Gupta

On Quality Improvement of a few Multi-Objective Combinatorial Optimization Problems with Hybridization Approach

Performance measurement is a complex issue in multi-objective combinatorial search. Not a single metric is superior in terms of all performance issues like coverage of solutions to the Pareto front, diversity across Pareto front, etc. Monitoring performance metrics on Pareto front is a step towards the improvement of quality of solutions. Having a set of test functions and metrics which could un-ambiguously monitor the performance of Multi-Objective Combinatorial Optimization (MOCO) problems is a challenging area of research. The scope of hybridization with EA approach in MOCO field is widely open due to small amount of research done with this approach.

We aim to devise a way to hybridize MOEA which improves quality of solutions for a set of MOCO problems and assess the performance across the solution sets. We aim to propose a set of performance metrics which can assess the performance of problem-specific MOEA solutions quantitatively.

On Quality Improvement of Bounded-Diameter MST Instances with Hybridization of Multi-Objective EA

We have recasted a few well-known heuristics, which are evolved for well-known Bounded Diameter (a.k.a Diameter Constraint) Minimum Spanning Tree (BDMST/DCMST) problem to a Bi-Objective Minimum Spanning Tree (BOMST) problem and then obtained Pareto fronts. A detailed analysis of Pareto fronts suggests that none of the heuristics provide superior solutions across the complete range of the diameter. In this work, we have used a Multi-Objective Evolutionary Algorithm (MOEA) approach to improve the Pareto front for BOMST, which in turn provides better solution for BDMST instances. We have considered edge-set encoding to represent MST and then applied recombination operators having strong heritability and mutation operators having negligible complexity to improve the solutions. The analysis of MOEA solutions confirms the improvement of Pareto front solutions across the complete range of the diameter over Pareto front solutions generated from individual heuristics.

Characterization of Graph Properties for Improved Pareto fronts using Heuristics and EA for Bi-Objective Graph Coloring Problem

Bi-Objective Graph Coloring Problem (BOGCP) is a generalized version in which the number of colors used to color the vertices of a graph and the corresponding penalty which incurs due to coloring the end-points of an edge with the same color are simultaneously minimized. In this work, we have analyzed the graph density, the interconnection between high degree nodes of a graph, the rank exponent of the standard benchmark input graph instances and observed that the characterization of graph instances affects the behavioral quality of the solution sets generated by existing heuristics across the entire range of the obtained Pareto fronts. We have used Multi-Objective Evolutionary Algorithm (MOEA) to obtain improved quality solution sets with the problem specific knowledge as well as with the embedded heuristics knowledge. To establish this fact for BOGCP, hybridization approach is used to construct recombination operators and mutation operators and it is observed from empirical results that the embedded problem specific knowledge in evolutionary operators helps to improve the quality of solution sets across the entire Pareto front; the nature of problem specific knowledge differentiates the quality of solution sets.

References

[1] K. Deb, Multi-Objective Optimization using Evolutionary Algorithms. John Wiley & Sons, Chichester, 2001.

[2] A. Singh and A.K. Gupta, “Improved heuristics for the bounded diameter minimum spanning tree problem”, Journal of Soft Computing, 11: 911-921, 2007.

[3] R. Kumar, 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.

[4] G. R. Raidl and B. A. Julstrom, “Greedy heuristics and an evolutionary algorithm for the bounded diameter minimum spanning tree problem”, 18th ACM Symposium on Applied Computing (SAC’03), 747-752, ACM, 2003.

[5] P. Galinier and J. K. Hao, “Hybrid evolutionary algorithms for graph coloring”, Journal of Combinatorial Optimization 3 (4): 379–397, 1999.

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, he has been a research scholar at the department of Computer Science and Engineering in IIT Kharagpur. His research interests are in the areas of Wireless Internet.*

Supervisor: Prof. Niloy Ganguly

Applications and solutions for next generation Wireless Internet

Wireless Internet provides anytime and anywhere connectivity to its users. Increasing numbers of users (approximately 2.3 billon users currently), continuous demand for better quality as well as the need for providing new innovative applications make research in the field of wireless network very challenging. There are lots of important areas where contributions need to be made – out of which we are concentrating on three problems: (a) Developing smart association control scheme and (b) Developing a collaborative download scheme whereby effective throughput of each individual mobile user can be enhanced. (c) Developing a content delivery system for wireless network which in effect can reduce overall traffic congestion.

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.

Content Distribution in Wireless Network

Due to the availability of cheap handheld devices and ubiquitous wireless connectivity, a huge demand for content has been noticed from wireless users. In year 2010, total wireless Internet traffic is 37% of overall Internet traffic and it has been predicted by 2015, it will cross 50%. This fact motivates to re-think for some new content delivery technique for wireless network which can reduce network traffic. To reduce the Internet traffic, one possible solution is to distribute the content utilizing some underutilized network resource (like WiFI in road network) for that local area.

Kamalesh Ghosh

Email: kghosh@cse.iitkgp.ernet.in
Joined the department in: July, 2009

***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

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.

Chester Rebeiro

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

***Chester Rebeiro** received a 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: Dr. Debdeep Mukhopadhyay

Design and Analysis of Secure Systems in the Presence of Side-Channel Leakages

In the last decade it has been shown that almost every secure system in use today is vulnerable to a class of cryptographic attacks known as side-channel attacks. These attacks glean secret information through leakages from power, timing, and electro-magnetic radiation of the device.

Preventing these attacks is difficult because the leakage not only depends on the cipher algorithm but also on the implementation and the execution platform. The counter-measures proposed in literature so far are ad-hoc and are either too difficult to implement or have large overheads. Moreover, most proposals only increase the complexity of the attack but do not prevent it. Theoretical analysis of side-channel attacks is critical in order to provide a fair evaluation of leaking crypto-systems. However, developing such an analysis is challenging due to the device and implementation specific nature of side-channel attacks.

The first step in formally analyzing side-channel attacks is to quantify the amount of information leaked from the implementation. Contemporary approaches abstract leakage from the physical devices by polynomial time functions. However this is known to correspond to more powerful leakages than what is actually observed in practice. An alternate approach is to approximate leakages by Hamming weight and distance models. Leakage, however, is a function of several parameters and the magnitude of leakage of each parameter may differ. For example, in software implementations of ciphers, leakage is influenced by numerous system specific parameters such as the branch prediction algorithm, hyperthreading, technology of the memory used, and the cache architecture. Hamming weight or distance models does not always apply in these cases. Depending on the cipher algorithm, its implementation, and form of side-channel attack, the leakage contribution of each parameter would vary. Therefore, in order to have an accurate measure of information leakage, it is important to pin-point the causes of leakage and quantify the amount of information leaked from each source. In my research with Dr. Debdeep Mukhopadhyay, we consider symmetric key ciphers implemented with look-up tables. In such implementations, the cache memory is the major source of leakage. We discovered that micro-architectural features in cache memories, such as non-blocking reads, out-of-order execution, parallelization, pipelining, and prefetching in memory accesses have a significant contribution in the leakage. Further we were able to mathematically quantify the leakage in commonly used prefetching algorithms. The analytical results, which were supported by experimentation, brought out interesting facts like the impact of the size, number of look-up tables and their relative placement on the information leakage. We are currently working on extending this analysis to the other micro-architectural features in the cache. This would give a comprehensive insight about information leakages in cache memories.

In the future, we plan to utilize the leakage models developed to construct a cipher provably resilient against side-channel attacks. The final objective is to implement the proposed cipher and then and then compare its side-channel resistance against against state-of-the-art ciphers like AES, CAMELLIA, and CLEFIA. The hope is the emergence of a new class of ciphers, more resistant against these lethal forms of attacks.

Satya Gautam Vadlamudi

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

Satya Gautam Vadlamudi *received a B.Tech. (Hons.) degree in computer science and engineering 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. He received the Pratibha Award from the Government of Andhra Pradesh (2004), the MCM Scholarship from the IIT Kharagpur (2004-2008), and the SAP Labs Doctoral Fellowship (2010-present) for his academic and research works. He is a graduate student member of IEEE.*

Supervisor: Prof. Partha Pratim Chakrabarti

Heuristic Search (Artificial Intelligence)

Heuristic search methods based on A* and beam search are widely used for solving several optimization and path planning problems. Due to large search spaces of the problems, anytime methods were developed for producing good quality solutions quickly. However, such anytime algorithms run out of memory when dealing with large sized search spaces.

We proposed a memory-bounded anytime heuristic search method called MAWA* (Memory-bounded Anytime Window A*) which works within the given amount of memory (should be at-least sufficient to hold all nodes of a single path from the start node to a goal node, which is usually very small), and still gives very good anytime performance. MAWA* is developed by intelligently combining AWA* (an anytime algorithm) and MA* (a memory-bounded algorithm). Also, it guarantees to terminate with an optimal solution.

Future directions include, developing better memory-bounded anytime algorithms, contract search (given a fixed amount of time, memory—get the best possible solution), etc.

References:

[1] P. P. Chakrabarti, S. Ghose, A. Acharya, and S. C. de Sarkar. 1989. Heuristic search in restricted memory (research note). *Artif. Intell.* 41, 2 (December 1989), 197-222.

[2] S. Aine, P. P. Chakrabarti, and R. Kumar. 2007. AWA*-a window constrained anytime heuristic search algorithm. In *Proceedings of the 20th international joint conference on Artifical intelligence* (IJCAI'07), 2250-2255.

Design and Validation of Dependable Systems

Embedded control systems are widely used in several domains such as aeronautic, aerospace, automotive, nuclear, medical, etc. The components often carry out highly safety-critical operations, which makes it extremely vital for them to be fault-tolerant. Logical faults such as ECU (Electronic Control Unit) failure, link failure were well studied.

We proposed the quality-fault model where we consider small amounts of shift, noise, and spikes on different signals of the electronic control system simultaneously (which can happen due to sensor faults, submicron chip technologies being used, etc.), and studied the behaviour of the system over a period of time for any unacceptable deviations from golden behaviour. We have proposed an efficient framework for finding the counterexamples (violating the given fault-tolerant requirements) which does static analysis on the network of abstract models of each component of the system.

Future directions include, developing more efficient and higher coverage fault-tolerance analysis methods/frameworks, and to explore reconfigurable architectures for achieving fault-tolerance at low cost.

Spatio-Temporal Data Mining

Spatio-temporal data mining of mobile user data such as the times series of locations of all individuals over a period of time is extremely useful for mining several interesting patterns. Groups, events, trajectories, hotspots, etc. have been mined in the past for using in social-network analysis, traffic modeling, crime investigation, etc.

We attempt to model novel queries which can be of use, such as, finding groups that cover maximum number of users, finding a mobility model that best fits a person, or a group, finding large sized gatherings, such as, a cricket match, or a wedding, etc. We aim to develop methods which can answer such hard queries quickly (seconds/few minutes), and which can scale over large number of users and large time periods.

Also, the mobility data itself is limitedly available due to privacy reasons, therefore, realistic mobility simulator is being developed for testing the mining methods.

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, in 2001 and 2004, respectively, and M.S. in Computer Science and Engineering from Indian Institute of Technology Kharagpur, in 2009. He is currently pursuing his Ph.D. in Computer Science and Engineering at Indian Institute of Technology Kharagpur. His research interests are in the area of algorithms for computer-aided design and testing of digital microfluidic biochips.*

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

Algorithms for Solution Preparation in Digital Microfluidic Biochips

Microfluidic-based biochips are recently emerged technologies and 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]. 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 techniques, microchip fabrication technology, etc. Typically, a biochip can implement one or more biochemical laboratory protocols or assays on a single chip that is a few square centimeters in size.

One category of microfluidic chips are continuous-flow microfluidic chips, where continuous liquid flow through micro-fabricated channels is manipulated with the help of micropumps, microvalves, etc. Another 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 sample/reagent fluids are manipulated on a substrate of two dimensional array of electrodes using electrical actuation (similar to 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/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.

To build a biochip efficiently, several associated combinatorial optimization techniques are applied. Recently, many CAD algorithms and techniques are being developed for both design and testing of DMF biochips [1]. Sample preprocessing and solution preparation for a biochip can be performed off-chip. However, that increases the overall biochemical assay time. Hence, 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.

A key challenge in design automation of DMF biochips is to carry out the dilution and mixing of fluids on-chip. We are developing algorithms for automated dilution and ratioed mixing of sample/reagent fluids. An existing algorithm for dilution/mixing of fluids minimizes the number of mixing steps to achieve the target concentrations [4]. However, in a bioprotocol, waste droplet handling is cumbersome and the number of waste reservoirs on-chip 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 bit-scanning (BS) method in prior work [4]. 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. Recently, we present another scheme for automatic solution preparation (referred as ratioed mixing algorithm or RMA) that can be applied for the solution preparation of more than two reagent fluids. It uses the (1:1)-mixing model at each mix-split step in the mixing tree to achieve the target ratio of component fluids. It has been shown that RMA is better layout-aware compared to the BS method [4].

References

[1] K. Chakrabarty and T. Xu, “Digital Microfluidic Biochips: Design and Optimization”, CRC Press, 2010.

[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] W. Thies et al., “Abstraction Layers for Scalable Microfluidic Biocomputing”. Natural Computing, Vol. 7, pp. 255–275, 2008.

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.*

Supervisors: Prof. Niloy Ganguly (IIT Kharagpur) and Dr. Monojit Choudhury (Microsoft Research India)

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] A. 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.

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

Power Aware Software

It has been observed that power reduction can be achieved at a higher degree at higher level i.e. at algorithmic level than at circuit and gate level. In past decades of CMOS technology dynamic power dissipation dominated the leakage power. As CMOS technology is reducing in dimensions leakage power is gradually becoming a challenge for power awareness. In addition to low power circuits, it is necessary for the system software like compilers and operating systems to be well equipped to achieve low power. Modern processors run on multiple voltage-frequecy pairs. Multicore processors have evolved keeping power and thermal management in mind. These have provided a room for designing low power software.

Compilers should deal with code optimization techniques to reduce power consumption of program. Early research on code optimization focused on time and space. Some of them reduce power consumption while others like software prefetching increases power consumption. In [1] we have proposed a scheme for power-aware software prefetching, where the software prefetching program trades performance for low power dissipation on an Xscale processor. We are working on compiler optimization techniques to minimize both dynamic power and leakage power consumption.

Operating Systems should care process, memory, I/O and file management to achieve low power. Achieving low power is challenging for real time systems as it may degrade performance. Power and thermal aware task scheduling for multi-processors/multi-core processors is an important area of our research. As, multi-processor task scheduling is an NP-complete problem. Sophisticated battery enabled systems require task scheduling techniques that will elongate the battery lifetime. We are also working on power-aware memory management techniques for garbage collection of Kilobyte Virtual Machine (KVM), the Java Virtual Machine (JVM) for Java enabled embedded systems.

There is a lot of scope in designing software for low power in different domains of computer science. Some of them are Database Management System and Computer Networks. As both of them is the most important part of today’s industry. Power aware database query optimization is an area of our concern. Network protocols right from data link layer to application layer should be power-aware because most of today’s hand-held devices provide networking facilities.

[1] S. Pyne, K. Ray, and A. Pal, Realization of Power Aware Software Prefetching as a Multi-Objective Optimization Problem', second IEEE International Conference on Computer and Communication Technology (ICCCT-2011), Allahabad, 2011.

.

Bibhas Ghoshal

Email: bibhas@cse.iitkgp.ernet.in
bibhas.ghoshal@gmail.com
Joined the department in: July 2008

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

Supervisor: Prof. Indranil Sengupta

Devising Improved Techniques for Testing Network-on-Chip Based Memory Systems

The increasing demand for fast and more features but low overall hardware cost of the product has led to growth in the demand for Systems-On-Chip (SOCs) with faster, better and smaller chips. As SOCs are packed with multiple cores, there is a steady increase in the amount of memory embedded into an SOC. The embedded memory content in SOCs have increased from one-tenth to more than three fourth of the chip area today and will continue to increase. Due to their high density, these embedded memories are more prone to manufacturing defects than other type of on-chip circuits and it is important to test them thoroughly to ensure quality of products shipped to customers. A lot of research is devoted in finding efficient testing methods for memory cores at reduced test time and power at minimum area overhead. With more and more memories embedded in circuits, accessibility becomes an issue in tester based methods making Built-in-Self Test (BIST) the solution of choice.

BIST reduces pin count at system level; requires no external test equipment; reduces development efforts; on-chip test pattern generation to provide higher controllability and observability; on-chip response analysis; tests are carried out at speed and test time gets reduced as number of cores can be tested parallely. However, unless properly implemented, main limitations of BIST lie in high power dissipation and area overhead, and increased testing time, all of which directly influence the cost and quality of manufacturing test. For memory BIST insertion, a tradeoff between area, test power and test time has to be considered.

My research work aims at devising a BIST based methodology for testing heterogeneous memory modules embedded in SOCs for optimized test time, area and power. The methodology includes designing a suitable Network-on-Chip (NOC) based test architecture to increase test parallelism by testing a number of cores simultaneously. This reduces the total test time of memory cores and reusing NOC as Test Access Mechanism (TAM) brings down the area overhead. Since test scheduling under power constraints is related to resource sharing mechanisms in Memory BIST architecture, the memory test methodology used in my research also includes development of a proper test schedule under power constraints that will get maximum usage of available resources.

The test of the complete system also requires testing of the NOC fabric. Through my research work, I intend to study the role played by the NOC communication fabric in the overall test time of the system and develop suitable BIST based techniques for evaluating the contribution of the communication fabric in the overall test time at optimized test power and area overhead.

Subhasish Dhal

Email: subhasis.rahul@gmail.com
Joined the department in: January 2010

Subhasish Dhal *received a B. Sc. degree in Computer Science (H) from Midnapore College under the affiliation of Vidyasagar University in 2002. He received an MCA degree from National Institute of Technology Durgapur in 2005 and an M. Tech degree in Computer Science from National Institute of Technology Rourkela in 2009. From august 2005 till august 2007, he worked in Asutosh college Kolkata as a lecturer. From august 2009 till December 2009, he worked in Institute of Engineering and Industrial Technology Durgapur as a lecturer. 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 RFID and Mobile Networks.*

Supervisor: Prof. Indranil Sengupta

RFID and Mobile Networks

RFID defeats bar code technology in respect to efficiency. Therefore, the objects are tagged with RFID devices which contains relevant information related to the object. This information may be secret and hence the access needs to be secure. Therefore, issues such as privacy, authentication, integrity, etc constitute some at the basic requirements for these devices. In this work, we focus on the authentication issue. Many proposals have been made by earlier researchers for the authentication of RFID devices. However, these proposals are based on the assumption that each object would be having single RFID tag. Use of multiple RFID tags in an object increases the detection rate in comparison to single tagged objects. Since more than one tag is associated with the same object, security issue like authentication needs to be looked at. The advantage of more resources (tags) can be utilized to enhance the security. Our goal is to utilize the extra resources and propose authentication schemes which are not only applicable to multi tag environment but also enhance the security.

Searching a particular tag from a pool of tags is known tag searching. This is another area we focus in this work. We may use any authentication scheme to search a desired tag. However, instead of authenticate every tag in the pool; we require authenticating only the desired tag. In multi tag environment, an object is attached with multiple numbers of tags.

Our motivation is to search an object attached with a particular tag from a pool of objects. We aim to propose a tag searching scheme in multi tag environment which is secure and efficient.

Identification of simultaneous existence of more than one related tags is tag binding. We require checking the authentication of relevant tags exists simultaneously. In multi tag environment, our motivation is to verify the simultaneous existence of objects attached with relevant tags. This is the third area we want to focus in this work. We aim to propose a tag binding scheme in multi tag environment.

Kunal Banerjee

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

***Kunal Banerjee** received his 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

Formal Verification of Embedded System Designs using FSMD models and their Extensions

Embedded systems are generally defined as special-purpose computer systems designed to perform one or more dedicated functions, usually with real-time constraints, and have computer hardware and software embedded as parts of a complete device or system, ranging from simple portable devices to sophisticated multiboard stationary systems. Because of the significance of the service they deliver and the increasing complexity of their designs, a lot of study has been invested to guarantee their behaviour. Testing fails to reach this goal due to the exponential number of input cases, which has led to an emphasis on formal verification to assure safety and appropriate functionality.

A finite state machine with datapath (FSMD) [1] is a universal specification model that can represent all hardware designs. A path-extension based method for checking the equivalence between two FSMDs, representing the original and the transformed behaviours, has been proposed in [2]. However, this method, can not verify code motion techniques that involve operations moved beyond loops. To alleviate this problem, we have developed a value propagation based method. The objective is accomplished by repeated propagation of the mismatched values to subsequent paths in an FSMD until the values match or the final path segments are traversed without finding a match. Checking loop invariance of the values being propagated beyond the loops has been underlined to play an important role. The proposed method is capable of handling control structure modification as well. The complexity analysis depicts identical worst case performance as that of a related earlier method of path extension which fails to handle code motion across loops. The method has been implemented and satisfactorily tested on the outputs of a basic block based scheduler, a path based scheduler and the high-level synthesis tool SPARK for some benchmark examples.

The next objective is to ensure the correctness of loop and arithmetic transformations in array-intensive programs. The array data dependence graphs (ADDGs) [3] are employed to achieve this goal. However, ADDGs suffer from the following shortcomings: single assignment form, non-uniform recurrences, no provision for specifying data-dependent index ranges and data-dependent control structures. The FSMD model is extended with arrays (FSMD+A) by us in order to overcome these deficiencies. Moreover, ADDGs, being able to capture only the data flow graphs involving arrays, have found applications mainly in the multi-media domain, whereas we aim at catering to a larger set of programs involving scalars and arrays that have undergone data as well as control transformations. Furthermore, the mappings from the input arrays to the output arrays for the ADDGs corresponding to the original and the transformed behaviours are constructed in isolation before performing equivalence checking between them. In contrast, for equivalence checking of two FSMD+A models, computation of such mappings proceeds hand-in-hand as paths in the two models are compared; the process reveals points of symmetry (and asymmetry) between the two behaviours. Hence, it is anticipated that in case of non-equivalence, the procedure involving FSMD+A models will report it much earlier than that of ADDGs, pin-pointing the regions where they mismatch and therefore be of more help for debugging purposes.

References

[1] D. D. Gajski, N. D. Dutt, A. C. Wu, and S. Y. Lin, “High-Level Synthesis: Introduction to Chip and System Design”, Kluwer Academic Publishers, 1992.

[2] C. Karfa, D. Sarkar, C. Mandal, and P. Kumar, “An equivalence-checking method for scheduling verification in high-level synthesis”, IEEE Transactions on CAD of ICS, 27:556 - 569, 2008.

[3] K. C. Shashidhar, M. Bruynooghe, F. Catthoor, and G. Janssens, “Functional equivalence checking for verification of algebraic transformations on array-intensive source code”, DATE, pp. 1310 - 1315, 2005.

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] R. Frost, K. I. Forster, and A. Deutsch. (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] W.D. Marslen-Wilson, L. K. Tyler, R. Waksler, and L. Older. (1994). Morphology and meaning in the English mental lexicon. *Psychological Review*, *101*, pp. 3–33.

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. He has published several research papers in various international conferences and journals including a Best Student Paper award in VLSI Design Conference (2010). He has also been awarded with the Microsoft Research (India) Ph.D. Fellowship in the year 2011.*

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

Formal Methods for Architectural Verification of Power Intent and Functional Reliability Analysis

The growing trend towards using component-based hierarchical design approach in system development requires addressing newer system engineering challenges. Based on an overall architectural plan, large designs are built hierarchically starting from the leaf-level components and thereby gradually constructing the overall system (bottom-up). During the development phase of such component-based hierarchical systems, the designers need to guarantee four critical aspects of design, namely – *correctness*, *timing and performance*, *power intent* and *functional reliability*. A significant amount of research has been conducted on the aspects such as *design intent verification* (to ensure correctness) and *timing/performance analysis* of component-based designs. Emerging paradigm, like design intent coverage, ensures that the component-level formal specifications together guarantee the end-to-end system requirements. Moreover, the component-based complex systems of modern days are usually very time-critical and require timing guarantees from components. So for system integration, the current research also focuses on introducing new notion of overall system timing layout by specifying the time-budgeting for its constituent components. However, meeting a stringent low-power budget by efficient global power management schemes and meeting the desired level of functional reliability for the overall system are emerging issues now-a-days. This work tries to address (from architectural point of view) the following two important aspects of component-based hierarchical system design: *power intent verification* and *functional reliability analysis*.

Recent research has indicated ways of using unified power format (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. Our work presents a verification framework that attempts to bridge the disconnect between high-level properties capturing the architectural power management strategy and the implementation of the power management control logic using low-level per-domain control signals. The novelty of the proposed framework is in demonstrating that the architectural power intent properties developed using high-level artifacts can be automatically translated into properties over low-level control sequences gleaned from UPF specifications of power domains, and that the resulting properties can be used to formally verify the global on-chip power management logic. The proposed translation uses a considerable amount of domain knowledge and is also not purely syntactic, because it requires formal extraction of timing information for the low-level control sequences.

Apart from the correctness, the completeness of the global power management logic also needs to be examined formally by the help of its global power state coverage. The architectural power intent of a design defines the intended global power states of a power-managed integrated circuit. Verification of the implementation of power management logic involves the task of checking whether only the intended power states are reached. Typically, the number of global power states reachable by the global power management strategy is significantly lesser than the possible number of global power states. In this work, we present a formal method for determining the set of reachable global power states in a power-managed design. Our approach demonstrates how this task can be further constrained as required by the verification engineer. In our work, we also developed a tool, called *POWER-TRUCTOR* which enables the proposed framework to guarantee the correctness and completeness of a global power manager.

In addition to this, the power consumption of complex System-On-Chips (SOCs) heavily depends on the application profile and the usage patterns. Customization of power management strategies based on application as well as user profile has great promise towards minimizing the power wastage. In our work, we present a methodology for computing the power usage at a higher level of granularity for analyzing the scope of the optimization of the power management strategies in a complex low-power SOC with respect to given workload profiles by customizing the boundaries of power domains. The proposed methodology computes the power usage for a given SOC and a given power domain partitioning. We have implemented the methodology using SystemVerilog platform as well as using JAVA platform and discuss the relative merits of each platform. We further demonstrate the effect of user profiles along with other major factors like task mapping etc. on the partitioning strategy.

Though the structural reliability of logic circuits as well as component-based embedded systems are well studied, however early analysis of the *functional reliability* of a component-based design is becoming important in high integrity systems. In our work, we present formal methods for determining whether a set of components having given reliability certificates for specific functional properties are adequate to guarantee desired end-to-end properties with specified reliability requirements. We also introduce the formal notion of *reliability gap* in component-based designs. This analysis opens the avenue for developing suitable procedures to bridge the functional reliability gap. In addition to this, there is a need to formally model the functional reliability of component-based systems.

Therefore, this work, as a whole, aims to enable formal methods for addressing two important aspects of component-based designs (from the architectural perspective), namely, power intent verification and functional reliability analysis.

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 14 international conference papers and journals in this domain.*

Supervisor: Prof. Indranil Sengupta and Prof. Santanu Chattopadhyay

Fault Diagnosis in Digital Systems

When a logic circuit fails a test, diagnosis is the process of narrowing down the possible locations of the defects. Fault diagnosis is extremely important to ramp up the manufacturing yield especially for 90 nm and below technologies where physical failure analysis machines become less successful due to reduced defect visibility by the smaller feature size and larger leakage currents. Diagnosis helps to reduce the product debug time as well. By reducing the candidate locations down to possibly only a few, subsequent physical failure analysis becomes much faster and easier when searching for the root causes of failure.

A failure can occur in a circuit due to the defects present in the logic circuit or in the scan chains. While many defects reside in the logic part of a chip, defects in scan chains are becoming more and more common, as typically 30%-50% logic gates impact the operation of scan chains in a scan environment. Logic diagnosis and scan chain diagnosis are the two main fields of diagnosis research. In this research work, we propose to consider both the problem of logic and scan chain diagnosis and attempt to find a suitable test methodology to assist in finding out the defects that possibly caused the circuit to fail.

The followings are the objective of the research work:

- **Multiple fault diagnosis in combinational logic circuit**

As stated earlier, with the ever-increasing complexity of VLSI circuits, multiple faults have now become a reality. Almost all the conventional fault diagnosis methods 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 propose to inject multiple faults simultaneously, and perform an effect-cause analysis to find the possible list of faults. 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 propose to use a Particle Swarm Optimization (PSO) algorithm for finding suitable solutions. Initially, a list of possible fault candidates will be found out by critical path tracing from each failing primary outputs (PO) and taking a union of them. Theoretically, our algorithm should be able to diagnose successfully up to any number of faults present in the circuit.

- **A diagnosability metric for test set selection**

The primary focus of a diagnosis algorithm is to accurately narrow down the list of suspected candidates. For that, all the diagnosis algorithms depend on the failure information produced by the tester. Some diagnosis algorithms also use the pass patterns to narrow down the list further. Overall, the backbone of any diagnosis algorithm is the test set in use. Thus, for any diagnosis algorithm, the effectiveness will depend on the test set in use. If the test set used is not good enough to distinguish between fault pairs, the diagnosis algorithm can never be able to distinguish between a good number of faults. This problem leads us to find a metric which can characterize test sets in terms of their diagnostic power. In literature, several methods have been proposed for assessment of the diagnostic power of a test set. Though the methods are accurate in nature, the bottleneck is the space and time complexity. In this work, we propose to find out a metric to describe diagnostic power of a test set. We call this metric, the diagnosibility of the test set for a given circuit.

- **Multiple chain failure in space response compaction environment**

Due to limited tester memory, test response compaction for large circuit has now become a necessity. Consider a design with approximately 500000 flip-flops. In addition, suppose that 30000 test patterns (e.g., stuck-at and transition fault test patterns) are used during scan test. The total test response data volume with traditional scan will be:

$$30000 * 500000 * 2 = 3.75 \text{ GB}.$$

Clearly, a tester cannot store this much information. So, test response compaction is mandatory for today's large VLSI design. We assume that a simple parity tree is used for compaction (space compactor) of test responses.

Failures on multiple scan chains are observed much more frequently than they were before. Note that with the space compactor, the number of scan chains is much larger than that of traditional scan designs, thus the probability of having multiple scan chain failures is even higher in a modern scan compression design than traditional scan design.

When space compactors are used, internal scan chains are not observed directly at the output channel. The reduced observability of internal scan chains and the interaction between them can adversely impact scan-based diagnosis. Thus, we propose to develop a suitable methodology for diagnosis of multiple chain failure in response compaction environment.

Parantapa Bhattacharya

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

Parantapa Bhattacharya *received a B.E. degree in Information Technology from Bengal Engineering and Science University, Shibpur in 2008, and a 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 & Engineering in IIT Kharagpur. His research interests are in the areas of Online Social Networks, Cloud Computing, and Computer Security.*

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

Malware in Social Networks

Social networks have lately evolved to become a major component of communication and sharing in the Internet. The social network companies are striving to ease this process of sharing, so as to get an early advantage. This has provided spammers and malware authors a unique platform to deliver their payloads in an efficient and intelligent manner.

Studies have shown that, click through rates in malicious messages are significantly higher in social networks, when compared to email [Grier et al., 2010]. A large percentage of accounts used for malicious activities in social networks are compromised legitimate accounts. The basic sharing nature of social networks, provide attackers with a golden opportunity to gather information and use social engineering as a tool to make their payload delivery efficient and targeted [Abraham and Chengalur-Smith, 2010]. This when combined with the fact, that many users accept links or social connections from unknown users [Ghosh et al., 2012], poses significant threat to existing defense mechanisms deployed to stop them. Attackers also use a large number of fake accounts, much of which is automated. For example, the Koobface botnet, coerces compromised users into solving CAPTCHAs to automate their fake account creation process [Thomas and Nicol, 2010].

The general technique to stop malicious messages is to identify them using blacklisting services and removing them. But this has been shown to be ineffective as blacklisting services run at a lag of few days during which most of the intended victims are exposed to the message anyway [Grier et al., 2010].

To combat this scenario, we are currently in the process of developing predictive, heuristic systems, which can answer the following questions, with the behavioral and usage information and in absence of a real time blacklisting service:

- Given an account, how likely is it that it has been compromised?
- Given an account, how likely is it that it is a fake account?
- Given a legitimate account, how susceptible is it to being compromised?
- Given a cascade of messages with an url, how likely is it that the url is malicious?

The basic idea is to devise meaningful feature vectors, using the social structure and user behavior, for developing supervised and unsupervised learning techniques for the said purposes. We are also working on creating fake honeypot accounts, for the purpose of letting them to be compromised. Once compromised, we can monitor the behavior of these accounts to better understand current trends in spamming and malware distribution.

References

[1] S. Abraham and I. Chengalur-Smith. An overview of social engineering malware: Trends, tactics, and implications. Technology in Society, 32(3):183–196, 2010.

[2] S. Ghosh, B. Viswanath, F. Kooti, N. Sharma, K. Gautam, F. Benevenuto, N. Ganguly, and K. Gummadi. Understanding and Combating Link Farming in the Twitter Social Network. In WWW ’12, 2012.

[3] C. Grier, K. Thomas, V. Paxson, and M. Zhang. @spam: The Underground on 140 Characters or Less. In CCS ’10, pages 27–37, 2010.

[4] K. Thomas and D. M. Nicol. The Koobface Botnet and the Rise of Social Malware. In MALWARE ’10, pages 63–70, 2010.

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

Multimedia Content Protection through Reversible Watermarking

The last couple of decades have seen rapid growth of research interest in the field of reversible watermarking of multimedia data. Primary goal of reversible watermarking is to maintain perfect integrity of the original content after watermark extraction. Such a feature is desirable when highly sensitive data is watermarked, e.g. in military, medical and legal imaging applications. Our goal is to analyze, implement and evaluate existing and new reversible watermarking algorithms.

The work done till now in the direction of achieving our goal, includes an extensive literature survey covering five classes of reversible watermarking algorithms based on five different principles of operation. We performed a simulation based study of such algorithms in a highly noisy environment which is often encountered in the specific application domain of reversible watermarking, viz. military image communication systems. We have investigated reversible watermarking in halftone images. Direct application of existing schemes to different color components of a halftone color image leads to considerable loss of perceptual quality of the image. We have proposed a reversible watermarking scheme for halftone color images which causes negligible degradation of the perceptual quality of the cover image. We have also developed a grayscale image reversible watermarking algorithm which utilizes coordinate logic operation based prediction of pixel values. We have compared the performance of our scheme with other reversible watermarking schemes, and results indicate that the cover image distortion produced by the proposed algorithm is lower than other state–of–the–art reversible watermarking algorithms.

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.

Mahesh 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 Vidyapeeth's College of Engineering, Mumbai, as a Lecturer. Since April 2000, he has been in Veermata Jijabai 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 Algorithm.*

Supervisor: Prof. Rajiv Kumar and Prof. Arobinda Gupta

Transition Sequence Exploration in UML Behavioral Models for Test Case Generation

Model based engineering and testing (MBT) have evolved as suitable candidates over the last decade to face challenge of complex industrial software. Model-based testing is software testing, in which test cases are generated in whole or in part from a model that describes some (usually functional) aspects of the system under test (SUT). Unified Modeling Language (UML) [3] comprises a suite of diagrams, to capture and model different aspects of the complex system. MBT can choose right type of diagram for right behavior to generate test suite. To improve the quality of the test suite coverage criteria should be maximized and the size of test suite should be minimized.

Evolutionary algorithms represent a class of stochastic search techniques and procedures based on the process of natural evolutions. They are characterized by an iterative procedure and can work in parallel on the number of potential solutions. The fundamental concept of EA is to evolve successive generations of increasingly better combinations of those parameters that significantly affect the overall performance of a design. EA has attracted the researchers to generate test data, and test scenarios [1], [2]. However, not all the UML behavioral models have been explored by the researchers to generate test cases.

The proposed architecture of test case generation using UML diagrams is shown in Figure 1. This architecture accepts the UML model as input and produces a set of test cases as output. Input model is fully furnished design model. It includes major design details which help test case generation.

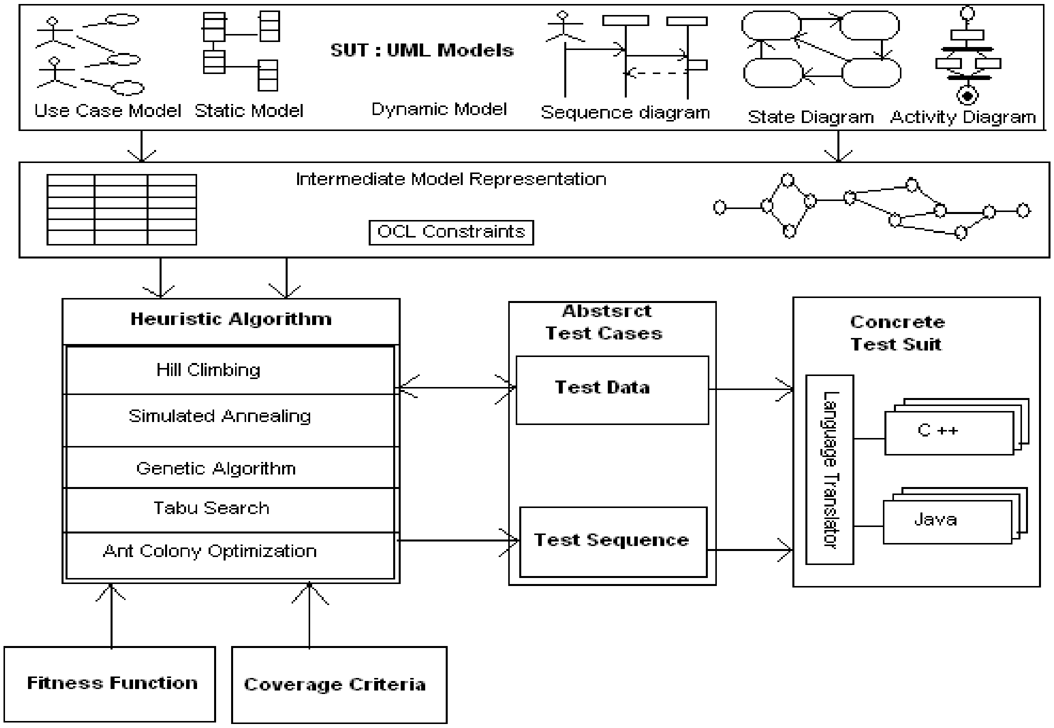


Figure 1: Proposed architecture for test suit generation from UML model using Heuristic Algorithms.

References

[1] P. McMinn. Search-based software test data generation: a survey. Software Test Verification and Reliability, 14:105-156, June 2004.

[2] S. Ali, L. Briand, H. Hemmati, and R. Panesar-Walawege. A Systematic Review of the Application and Empirical Investigation of Search-Based Test Case Generation. IEEE Transactions on Software Engineering, 36(6):742-762, 2010.

[3] OMG. UML superstructure v2.1.2, November 2007. http://www.omg.org/spec/UML/2.1.2/.

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

Information Management in Unstructured and Challenged Networks

Search and dissemination of information – are one of the most frequently used services in the Internet. The same is true for any general purpose distributed system. The underlying algorithms for these services try to cover as much area of the network as possible in lesser time. In unstructured networks, due to lack of any relationship between the placement of information and the node identities, the task of covering more area in lesser time becomes highly challenging. Unfortunately, because of its inherent resilience against many other 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 maximization of coverage under resource constraints – becomes a significant issue in these areas.

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 problem space. 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. However, 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 Dr. Subrata Nandi 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.

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. Using results from analytical studies done in the past on K-random walk based algorithm, we identify that redundant node visits by the message packets (modeled as walkers) quickly increases with increase in the concentration of the walkers. Based on this postulate, we design a very simple distributed algorithm which dynamically estimates the concentration of the walkers and thereby carefully proliferates walkers in sparse regions. We use extensive computer simulations to test our algorithm whereby we find it to be far better than the existing best known algorithms.

This research also focuses on the information management services for the delay/disruption tolerant networks (DTN) where the problems are more difficult to solve because of the lack of end-to-end path between two nodes. For this reason, these networks are also termed as challenged networks. Frequent network partitions caused by insufficient node (human agents) density coupled with mobility of the nodes, is one of the prime problems in DTN. In order to circumvent this problem, these networks are often augmented with message buffers, commonly known as relay points, which help in effectively increasing the contact frequency in between the mobile agents in the network. Different strategies for disseminating a given piece of information can be designed exploiting the infrastructure of these buffers and consequently boost up the achieved coverage (i.e., the number of distinct nodes to which the message could be delivered). The time duration for which a piece of information is stored continuously in a particular buffer, becomes a significant parameter in the process. The achieved coverage as well as different overheads such as number of redundant message copies, consumption of battery power, CPU processing - are directly related to this buffer time. Store-carry-forward paradigm as well as complex human mobility model where the next node to be visited is selected preferentially, make it very difficult to use traditional mathematical tools to calculate the coverage in the dissemination process for a given buffer time. Hence, calculation of the optimal buffer time to achieve a desired coverage for a specified cost also becomes hard.

We first identify that the dynamics of this information dissemination process in DTN matches with the evolution process of a bipartite network constructed appropriately from the DTN. Using the rich analytical backbone of this bipartite network, we also develop a mathematical formula for calculating the coverage as well as show a possible direction to derive the optimal buffer time. We do these analyses under simplifying assumptions and later show the effect of relaxing these assumptions.

Sandip Karmakar

Email: sandip1kk@gmail.com
Joined the department in: December, 2010

***Sandip Karmakar** received his B.E. degree in Computer Science & Technology from Bengal Engineering and Science University, Howrah, WB in 2004. After receiving his B.E., he worked in Software Industries from June 2004 to December 2007. Since May 2008, he has been a research scholar in the department of Computer Science & Engineering in IIT Kharagpur. He received an M.S. (by Research) from CSE of IIT Kharagpur in Oct 2010. Since Dec 2010 he is working towards his PhD degree in the same department. His research interests are in the areas of Cryptography, Cellular Automata and VLSI.*

Supervisor: Prof. Dipanwita Roy Chowdhury

Cryptanalysis and Design of Stream Ciphers

My broad area of research is *cryptanalysis and design of stream ciphers*. We are mainly working in *scan-based side channel attacks* and *fault attacks* on stream ciphers, both of which are *side channel attacks*. Another area of cryptanalysis we are currently working on is the algebraic cryptanalysis method *cube attack*. Our research also involves *design of cellular automata based stream ciphers*.

Scan chain based attacks are a kind of side channel attack, which targets one of the most important features of today's hardware - the test circuitry. Design for Testability (DFT) is a design technique that adds certain testability features to a hardware design. On the other hand, this very feature opens up a side channel for cryptanalysis, rendering crypto-devices vulnerable to scan-based attack. We have shown that the eStream ciphers, Trivium and Grain-128 can be analyzed using scan based side channel attack in a few minutes. We have proposed a more generalized approach which may break any cryptographic algorithm through scan chain interface in not more than few minutes and demonstrated it on hardware based eStream ciphers, Trivium, Grain-128, MICKEY 2-128. We also proposed a countermeasure to prevent such kind of attacks on stream ciphers.

Fault attacks are one of the most efficient form of side channel attack against implementations of cryptographic algorithms. In this attack, faults are injected during cipher operations. The attacker then analyzes the fault free and faulty cipher-texts or key-streams to deduce partial or full value of the secret key. The literature shows that both the block ciphers and stream ciphers are analyzable using fault attack. We have shown that the eStream cipher Grain-128 can be attacked by inducing faults in the NFSR. The attack requires about 56 fault injections for NFSR and a computational complexity of about 2^{21}, hence, it can be performed practically. Currently, we are working on multi-bit fault attacks on eStream ciphers and prevention of such kind of attacks.

Cube attack was introduced by Ita Dinur and Adi Shamir in Eurocrypt 2009. It is a kind of high order differential attack. The main challenge in this kind of attacks is finding cubes. We have proposed a heuristic to find cubes which was successfully applied to a simplified version of Trivium in less than 5 hours. Currently we are working on improving the existing results on Trivium using cube attacks and trying to apply cube attacks on other ciphers. Another area of cube attack that we are working on is the hardware implementation of cube attack, cube testers and dynamic cube attack. The hardware is to be designed to perform the mentioned attacks, distinguishers on state of the art stream ciphers.

The final area of our current research is the design of stream ciphers. During our previous research on design of stream ciphers using Cellular Automata we identified certain hybrid CA configurations that are cryptographically suitable. In the ongoing research, we have proposed cryptographic stream ciphers using the identified Cellular Automata configurations. We are now working on design of a stream cipher using Cellular Automata which is based on the hybrid cellular automata and provides high speed, achieves low power and cryptographic efficiency.

Joy Chandra Mukherjee

Email: mjoy1982@gmail.com
Joined the department in: July 2011

***Joy Chandra Mukherjee** received a B.Tech. degree in Computer Science and Engineering from Bengal Institute of Technology, Kolkata in 2004. From November 2004 till September 2007, he worked in CTS, Kolkata as an Associate. Since October 2007 to October 2008, he worked as an Assistant Systems Engineer in TCS, Kolkata. He received an M.Tech. degree in Computer Science and Engineering from Indian Institute of Technology, Kharagpur in 2011. Since July 2011, 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 Distributed Algorithms.*

Supervisors: Prof. Arobinda Gupta

Distributed Event-based System

Event-based system design decouples system components through events. An event is any occurrence of a happening of interest, i.e., a state change in some component of a system. The components of a system communicate by generating and receiving event notifications. An event notification service or publish/subscribe middleware mediates between the components of an event-based system and conveys notifications from producers (or publishers) to consumers (or subscribers) that have registered their interest with a previously issued subscription. The advantage of such a design is that neither the published notifications nor the subscriptions are directed toward specific components. The notification service decouples the components so that producers are unaware of any consumers and consumers rely only on the information published, but not on where or by whom it is published. This facilitates easy integration of autonomous, heterogeneous components into complex systems that are easy to evolve and scale.

Event-based computing applications may involve vast numbers of sensing devices. These devices can potentially produce an enormous volume of raw sensor data. This sensor data has the potential to overwhelm both the resources available in large-scale event-based computing systems, and the ability of users to respond to it. In particular, raw sensor data tends to be very low-level, and must be further processed, analyzed, and transformed into higher-level information in order for it to be meaningful to applications and users.

Large numbers of users may be interested in the data produced by sensors, and scaling up the task of reliably delivering relevant sensor information to users is a challenging one. Failures of sensing devices may be common, and disseminating data to users is subject to the difficulties inherent to all large scale distributed systems, such as node failures and network partitions. We are interested in the design of such large scale event detection systems and their applications.

Sudakshina Dutta
Email: sudakshina@cse.iitkgp.ernet.in
Joined the department: July 2011

Sudakshina Dutta *received a B.E. degree in Information Technology from Jadavpur University in 2007. From July 2007 to June 2009, she worked as Member of Technical Staff in Interra Systems India Pvt. Ltd. She joined Department of Computer Science and Engineering of Indian Institute of Technology Kharagpur and received an M. Tech degree in July, 2011. Since then, she has been a research scholar in this department and her research interest includes Formal Verification of Concurrent Systems.*

Supervisors: Prof. Dipankar Sarkar

Formal Verification of Concurrent Systems

Formal Verification is the act of proving or disproving the correctness of the algorithm in any software or hardware systems. Different aspects of modeling concurrent systems include simple case in which processes run completely autonomously to the more realistic setting where processes communicate in some way. Seemingly innocuous small concurrent programs have been known to exhibit completely unanticipated behavior that, in some cases, may lead to crashes in the critical systems. This is why different verification techniques for exhaustively checking are really required for correct execution of the concurrent systems.

Tripti Swarnkar

Email: swarnkar.tripti@gmail.com
Joined the department in: July 2011

Tripti Swarnkar *received an MCA degree from National Institute of Technology, Raipur(CG) in 1998, and an M.Tech. degree in Computer Science from Utkal University, Bhubaneshwar(Odisa) in 2005.At present working as an Associate Professor at Institute of Technical Education and Research ,Shiksa 'O' Anusandhan University, Bhubaneshwar Odisa. Since July 2011, she has been a research scholar in the department of Computer Science & Engineering in IIT Kharagpur. Her research interests are in the areas of Bioinformatics and Machine Learning.*

Supervisor: Prof. Pabitra Mitra

Gene Selection for Biological Data

Although the human genome sequencing project is almost over, the analysis has just begun. A DNA microarray can track the expression levels of thousands of genes simultaneously and can reveal large amount of data about the inner life of a cell. The challenge is to evaluate these huge data streams and extract useful information.

Given a series of microarray experiments for a specific tissue under different conditions our aim is to find genes which are more informative or are signature genes that robustly distinguish different classes. The problem is to select features, here our genes are our features that have the biggest impact on describing the results and to drop the features with little or no effect.

Noisy or irrelevant attributes make the classification or clustering task more complicated, as they can contain random correlation. Our aim is to filter out these features. When class labels of the data are assailable we use supervised feature selection, otherwise unsupervised feature selection is appropriate. Recently unsupervised feature selection has attracted a lot of attention especially in bioinformatics and text mining. In our work we are trying to work with an unsupervised feature clustering technique which does not need the class label information in the dataset and is thus suitable for both supervised and unsupervised learning.

Conventional methods of feature selection involve evaluating different feature subsets using some index and selecting the best among them. Unsupervised feature selection algorithms belong to the field of unsupervised learning. These algorithms are quite different from the major bulk of feature selection studies that are based on supervised methods, and compared to the latter are relatively over looked. Unsupervised studies, unaided by objective functions, may be more difficult to carry out, on the other side they convey several important theoretical advantages: they are unbiased, by neither the experimental expert nor by the data analyst. Without prior knowledge of class labels it performs well and reduces the risk of over fitting. The downside is that it relies on some mathematical principle. The method does not need any search space and, therefore, is fast. These unsupervised FS methods being selected for study can be broadly categorized in two categories. First category involve maximizing clustering performance, using some measuring index and second considers selection of features based on feature dependency and relevance.

We are working with clustering algorithm for feature selection, being focused on minimizing the information loss incurred in the process of feature reduction, as well as the minimization of the redundancy present in the reduced feature subset. Since our method does not rely on the class label information in the dataset, it should works for both supervised and unsupervised learning.

Sanjoy Pratihar

Email: sanjoy.pratihar@gmail.com
Joined the department in: July 2011

***Sanjoy Pratihar** received his BTech in Computer Science and Engineering from North Eastern Hill University, Shilong, India, and received his ME in Computer Science and Engineering from Bengal Engineering and Science University, Shibpur, India. Currently he is a PhD scholar in the Department of Computer Science and Engineering, Indian Institute of Technology, Kharagpur, India. His research interests include digital geometry, document image processing, graphics analysis, and intelligent human-computer interaction. He has served as a lecturer in the Department of Computer Science and Engineering, University Institute of Technology, The University of Burdwan, Burdwan, India. He has published 7 research papers in edited volumes and refereed conference proceedings.*

Supervisor: Prof. Partha Bhowmick

On Some Digital-geometric Applications of Farey Sequence

Background In the year 1816, John Farey invented an amazing procedure to generate proper fractions lying in the interval [0, 1], called the *Farey sequence* [9]. It remained unattended and unexplored for almost a century until the beginning of last century. And in recent times, with the emergence of various algorithms in the digital/discrete space, several interesting works have come up related with the Farey sequence [1-4].

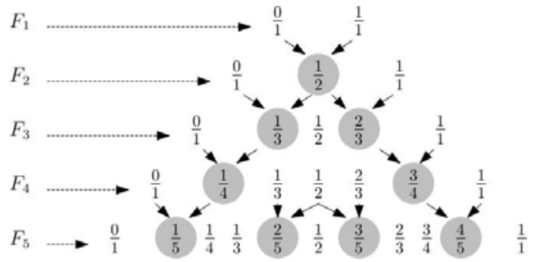
Our Idea of Augmented Farey Table The Farey sequence of order *n* is the sequence of simple/irreducible, proper, positive fractions with denominators up to *n*, arranged in increasing order (Fig. 1). The concept is well-known in *theory of fractions* [3, 9], but from the algorithmic point of view, very limited work has been done so far. In our work, we have augmented a Farey sequence with compound fractions, improper fractions, and negative fractions, which do not find any place in the original sequence. With all these *fraction ranks*, we build the *Augmented Farey Table (AFT).* We have used the AFT for several interesting applications, as mentioned below.

Polygonal approximation An efficient boundary representation of an object in the digital plane is done through polygonal approximation [10-12]. During approximation, "reasonably collinear" straight edges are successively merged. The collinearity is tested by edge slope, which corresponds to AFT rank. If the *rank difference* of two edges is less than a prescribed tolerance, then the two edges are merged into a single edge in an iterative manner. With the idea of *exponential averaging*, the AFT has been used by us for polygonal approximation in gray-scale images without any edge detection and thinning [5].

Shape Representation If all the internal angles are written in order for a polygon, we get an idea about its shape. As a novel alternative, we have used the sequence of rank differences corresponding to adjacent edges. This has subsequently been used for shape decomposition [6], shape matching [8], etc.

Vectorization of Thick Digital Lines Vectorization of a digital object provides a succinct, space-efficient, and useful representation for several applications in computer graphics and image analysis. As a fast and efficient vectorization of digitized engineering drawings, we have used AFT for geometric analysis and refinement [7].

Conclusion Usage of AFT enables all our algorithms to be devoid of floating-point operations, thus saving a significant amount of runtime. The notion of AFT also puts forward some important theoretical issues, such as compressing an AFT, as its size is quadratic with the order of Farey sequence.


F_1 $\frac{0}{1}$ $\frac{1}{1}$
F_2 $\frac{0}{1}$ $\frac{1}{2}$ $\frac{1}{1}$
F_3 $\frac{0}{1}$ $\frac{1}{3}$ $\frac{1}{2}$ $\frac{2}{3}$ $\frac{1}{1}$
F_4 $\frac{0}{1}$ $\frac{1}{4}$ $\frac{1}{3}$ $\frac{1}{2}$ $\frac{2}{3}$ $\frac{3}{4}$ $\frac{1}{1}$
F_5 $\frac{0}{1}$ $\frac{1}{5}$ $\frac{1}{4}$ $\frac{1}{3}$ $\frac{2}{5}$ $\frac{1}{2}$ $\frac{3}{5}$ $\frac{2}{3}$ $\frac{3}{4}$ $\frac{4}{5}$ $\frac{1}{1}$

		Denominator								
		-4	-3	-2	-1	0	1	2	3	4
N	4	19	18	16	14	13	12	10	8	7
u	3	20	19	17	15	13	11	9	7	6
m	2	22	21	19	16	13	10	7	5	4
e	1	24	23	22	19	13	7	4	3	2
r	0	25	25	25	25	–	1	1	1	1
a	-1	26	27	28	31	37	43	46	47	48
t	-2	28	29	31	34	37	40	43	45	46
o	-3	30	31	33	35	37	39	41	43	44
r	-4	31	32	34	36	37	38	40	42	43

Figure 1. Left: Farey sequences of orders 1 to 5. Right: AFT of order 4.

References

[1] G.H. Hardy and E.M. Wright , An introduction to the theory of numbers. Oxford University Press, New York, 1968.

[2] E.H. Neville., The farey series of order 1025. Cambridge University Press, 1950.

[3] C. E. Patrascu and M.Patrascu, Computing order statistics in the Farey sequence. ANTS 2004, pages 358-366.

[4] J. Pawlewicz, Order statistics in the farey sequences in sublinear time and counting primitive lattice points in polygons, Algorithmica, pages 55(2): 271-282, 2009.

[5] S. Pratihar and P. Bhowmick, A thinning-free algorithm for straight edge detection in a gray-scale image. In Proc. 7th Intl. Conf. on Advances in Pattern Recognition (ICAPR), pages 341-344. IEEE CS Press, 2009.

[6] S. Pratihar and P. Bhowmick, Shape decomposition using farey sequence and saddle points. In Proc. ICVGIP-2010, pages 77-84. ACM, 2010.

[7] S. Pratihar and P. Bhowmick, Vectorization of thick digital lines using Farey sequence and geometric refinement. In Proc. ICVGIP-2010, pages 518-525. ACM, 2010.

[8] S. Pratihar and P. Bhowmick, On applying the Farey sequence for shape representation in z2. In Speech, Image and Language Processing for Human Computer Interaction- Multi-modal Advancements (Accepted), IGI Global, 2011.

[9] D. Knuth R. Graham and O. Potashnik, In Concrete Mathematics. Addison-Wesley, 1994.

[10] P.L. Rosin, Techniques for assessing polygonal approximation of curves. In IEEE Trans. PAMI, pages 19(6): 659-666, 1997.

[11] M. Schroeder. Fractions: Continued, Egyptian and Farey (chapter 5), number theory in sc. and communication. Springer Series in Information Sciences, vol.7, 2006.

[12] P. Bhowmick and B. B. Bhattacharya. Fast Polygonal Approximation of Digital Curves Using Relaxed Straightness Properties, *IEEE Transactions Pattern Analysis and Machine Intelligence (TPAMI)*, Vol. 29, No. 9, pp. 1590-1602, 2007.

Tanmoy Chakraborty

Email: its_tanmoy@yahoo.co.in
Joined the department in: December 2011

***Tanmoy Chakraborty** received B.Tech degree in Computer Science and Engineering from Kalyani Government Engineering College, Kalyani, Nadia (affiliated to West Bengal University of Technology,Kolkata), in 2009 and M.E degree in Computer Science and Engineering from Jadavpur University, Kolkata in 2011. Since December 2011, he is a research scholar in the department of Computer Science & Engineering in IIT Kharagpur. His research interests are in the areas of Complex Networking, Graph Theory and Natural Language Processing.*

Supervisors: Dr. Animesh Mukherjee and Dr. Niloy Ganguly

Community Identification in Large Networks

Many complex systems like Social Networks, technological networks, biological networks share some distinctive statistical properties. One of them is *Community structure,* the division of network nodes into groups within which the network connections are dense, but among which they are sparser. The ability to find and analyze such groups can provide invaluable help in understanding and visualizing the structure of networks.

Though the traditional approaches in community detection based either on agglomerative, divisive or spectral analysis have been refined using new metrics like edge betweenness, outwardness, edge clustering-coefficient; new research challenges arise due to the intrinsic dynamicity of nodes and links. Some of them include detection of overlapping communities (nodes with equal involvement of two or more communities), static communities (nodes that are always allocated to the same community independent of the perturbations to the input parameters), mobility of nodes across communities over the time varying environment, detecting communities independently under the assumptions of input parameters, studies of different statistical behaviors of modules.

We have started our experiments of community detection using traditional algorithms. Besides, we are using citation networks from Computer Science domain to identify the impact of different fields (e.g. Algorithms, AI, Machine learning) and their influences towards creating new interdisciplinary fields (e.g. Bioinformatics, Hypertext). Our preliminary experimental results show some of the fields have high inwardness (e.g. Theoretical Computer Science) and few of them (e.g. Digital Library) have high outwardness vis-a-vis their inwardness. We will try to explore this idea over other domains like Physics, Biology, and evaluate the cross-linking behavior across the domains. We have plans to reconstruct the equation of impact and try to relate it with Google PageRank concept. Moreover, we will study the growth of impact of different fields over the time period and try to draw some interesting conclusions by the importance of emerging fields in sociological aspects like fund relies, shifting of field interest etc.

Anupam Mandal
Email: amandal@cse.iitkgp.ernet.in
Joined the department in: December 2011

***Anupam Mandal** received his B.E. and M.S. degree in Computer Science and Engineering from National Institute of Technology, Durgapur and Indian Institute of Technology, Madras respectively. He is currently a scientist at Center for Artificial Intelligence and Robotics, Bangalore. Since December 2011, he has joined the department of Computer Science & Engineering in IIT Kharagpur as a sponsored research scholar. His research interests are in the area of speech recognition and VoIP technologies.*

Supervisor: Dr. Pabitra Mitra

Keyword spotting in speech

My current work is on spotting keywords in continuous speech, a sub-area of continuous speech recognition. I am focusing on template-based approaches to keyword spotting that require lesser training data and may perform robustly in presence of noise and channel based degradations. As these methods involve matching of sound instances present in an utterance without any prior assumption of the underlying language, they may also work well for multilingual speech. My research is targeted towards novel methods of speech template representation and matching.

Durga Prasad Sahoo

Email: dpsahoo.cs@gmail.com
Joined the department in: December 2011

***Durga Prasad Sahoo** received B.Sc. degree in Computer Science from Ramakrishna Mission Residential College, University of Calcutta, Kolkata in 2007; M.Sc. degree in Computer and Information Science from University of Calcutta, Kolkata in 2009 and M.Tech. degree in Computer Science from University of Calcutta, Kolkata in 2011. From August 2011 till December 20011, he worked in Asutosh College, Kolkata, as a guest lecturer. Since December 2011, he has been a research scholar in the department of Computer Science & Engineering in IIT Kharagpur. His research interests are in the areas of Algorithm design, Graph Theory and Hardware Security.*

Supervisors: Dr. Rajat Subhra Chakraborty and Dr. Debdeep Mukhopadhyay

Machine Learning based Model-building Attacks on Physically Unclonable Functions

Counterfeiting of hardware devices and its impact on economy has become a big concern to modern society. The most well-known aspect of counterfeiting is product cloning. In order to deal with this aspect of counterfeiting, a secret unclonable identifier is required. The idea of using intrinsic random physical features to identify objects has led to the development of the concept of *Physically Unclonable Function* (PUF). The fact that PUFs are unclonable implies that they can be used for anti-counterfeiting purposes. When PUFs are used for the detection of the authenticity of a product, a physical property of the PUF is measured, translated into a bit string and verified. The physical unclonability of PUFs prevents building of a similar physical structure that upon interrogation produces a similar bit string that would pass the verification test as the original one.

However, recent studies on PUFs have challenged claims of unclonability by demonstrating that the behavior of PUFs, especially those implemented as solid-state electronic circuits, can be modeled by using machine learning techniques such as *logistic regression*, *perceptron learning*, *support vector machine*, etc.. Most common type of PUFs those are candidate for machine learning based attack are *Ring-Oscillator PUFs* and *Arbiter PUFs*. As a first step of research, we have applied above mentioned machine learning algorithms on challenge-response pairs of ring oscillator PUFs that is implemented on Altera FPGA board.

Tanwi Mallick

Email: tanwireachesu@cse.iitkgp.ernet.in
Joined the department in: December 2011

Tanwi Mallick *received a B.Tech degree in Computer Science from Jalpaiguri Govt. Engineering College in 2008 and an M.Tech degree in Computer Science from NIT Durgapur in 2009. From July 2010 to December 2011, she worked at DIATM, an engineering college of West Bengal. Since December 2011, she has been a research scholar in the department of Computer Science & Engineering at IIT Kharagpur. Her research interests are in the areas of Image Processing.*

Supervisor: Prof. Partha Pratim Das and Prof. Arun Kumar Majumdar

Analysis and Interpretation of Motion Information for Bio-Medical Applications

Processing of image and video is becoming popular for designing smart environments and automating health care services. Image segmentation and video object tracking are two of the most challenging problems that are faced while designing such environments. A segmentation algorithm that performs satisfactorily for a class of images may not perform as well for other class of images and hence selection of a suitable segmentation scheme is important. Similarly accuracy in tracking becomes important for building applications that rely on detecting and tracking of video objects. Image segmentation, application specific feature extraction and object tracking can help in analysis of video recordings of infant neurological examinations also known as Hammersmith Infant Neurological Examination (HINE). These set of examinations are used for assessing neurological development of infants between 2 and 24 months of age. Owing to the complex nature of test environment, handling of occlusion and noise is a challenging task. A set of machine learning techniques have been used to handle occlusion and noise and finally compute a score for Hammersmith tests.

MS Scholars

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

***Biswanath Barik** received a 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 & 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.

Debjit Pal

Email: debjit@cse.iitkgp.ernet.in
Joined the department in: January 2009

***Debjit Pal** received his B.E. degree in Electronics and Tele-Communication Engineering from Jadavpur University, 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 Simulation based Verification of AMS Circuits and Equivalence Checking of Transition Systems annotated with Predicates over real variables.*

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

Verification and Equivalence Checking of Analog-Mixed Signal Systems

Simulation has been the main technology for validating the integration of heterogeneous collection of analog/mixed signal (AMS) components into an integrated circuit. Given the increasing complexity of modern day AMS circuits with Big D (large digital components) Small A (small analog components) concept and lack of growth of mixed-signal simulation speed, it is becoming almost infeasible to verify the integrated circuits by checking simulation trace manually. One way to overcome difficulties in integrating system is to replace some of the most computation intensive components with their behavioral models that capture the design intent of the original component and verify the design. Also checking manually whether a property holds on the simulation trace of a circuit is infeasible as Mixed Signal simulation trace is continuum in terms of both time and value. To alleviate this issue, some intelligent property checker monitors may be developed which can automatically check the simulation trace and report satisfaction / violation of the properties. Finally, we can go further to check the equivalence of the digital controller components of the AMS circuits which have not only atomic propositions as their inputs but the truth / false of predicates over real variables can control their behavior. The focus of this research is to explore these directions.

Power Management Units (PMUs) [1, 2] are large integrated circuits consisting of many predesigned mixed-signal components. PMU integration poses a serious verification problem considering the size of the integrated circuit and the complexity of analog simulation. In this article we present an approach for automatic generation of behavioral models for PMU components from top-down skeleton models, fitted with parameter values estimated by bottom-up parameter extraction algorithms. It is shown that replacing PMU components with these auto-generated hybrid automata-based abstract behavioral models enables significant simulation speedup (> 20X on our industrial test cases) and helps in early detection of integration errors. The work also justifies the level of accuracy in our models with respect to the goal of verifying integrated PMUs. The approach presented in this work is implemented in the form of a tool suite called Chassis.

Our next objective is to develop a set of property checkers which can report satisfaction / violation of the properties automatically by examining simulation trace of a circuit. The development and use of assertions in the Analog and Mixed-signal (AMS) domain is a subject which has attracted significant attention lately from the verification community. Recent studies have suggested that natural extensions of assertion languages (like PSL and SVA) into the AMS domain are not expressive enough to capture many AMS behaviors, and that a library of auxiliary AMS functions are needed along with the assertion language. The integration of auxiliary functions with the core fabric of a temporal logic is non-trivial and can be challenging for a verification engineer. In this work, we propose a purely library-based verification approach, where libraries for checking elementary properties can be naturally connected with libraries for auxiliary functions to monitor complex AMS behaviors. We study the modeling of behaviors with the proposed library [1,3], and outline the main challenges and their solutions towards implementing the verification library over commercial AMS simulators.

Our final objective is to devise a fully symbolic methodology for checking simulation relation between the implementation and specification of the controllers of the hybrid systems like plant controller, AMS Circuit digital controller and controllers of other reactive systems. We are currently investigating this problem to propose a feasible solution.

References

[1] R. Mukhopadhyay, S.K. Panda, P. Dasgupta, J. Gough : Instrumenting AMS Assertion Verification on Commercial Platforms. In ACM Transactions Design Automation Electronic Systems (TODAES) vol. 14(2) (2009)

[2] A. Ain, D. Pal, P. Dasgupta, S. Mukhopadhyay, R. Mukhopadhyay and J. Gough, Chassis: A Platform for Verifying PMU Integration Using Auto-generated Behavioral Models, In ACM Transactions on Design Automation of Electronic Systems (TODAES), Article-33, Volume-16, Issue-3, June 2011.

[3] D. Pal, P. Dasgupta and S. Mukhopadhyay, A Library for Passive Online Verification of Analog and Mixed-Signal Circuits, In the proceedings of IEEE VLSI Design Conference, pp. 364-369, January 2012.

Anup Kumar Bhattacharya

Email: bhattacharya.anup@gmail.com
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.*

Supervisor: Prof. Dipanwita Roychaudhury & Dr. Abhijit Das

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.

Satrajit Ghosh

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

***Satrajit Ghosh** received a B.Sc (Honours) degree in Physics from Scottish Church College, University of Calcutta, Kolkata in 2005, and a B.Tech. degree in Computer Science from department of Computer Science & Engineering, University of Calcutta, Kolkata in 2008. Since June 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 Cryptology, Quantum computation, Algorithm design.*

Supervisor: Dr. Abhijit Das

Algebraic Attacks 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 (2^8) 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 XL_SGE 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.

Praloy Kumar Biswas

Email: praloy.slg@gmail.com
Joined the department in: July 2009

***Praloy Kumar Biswas** received a B.E. degree in Computer Science and Engineering from Jadavpur University, Kolkata in 2005, and presently doing his MS in Cryptography at Indian Institute of Technology, Kharagpur. After completing his B.E., he joined Wipro Technologies, Hyderabad as Project Engineer and worked there from September'2005 to November'2006. Then he switched the domain to EDA and worked at Atrenta Indian Pvt. Ltd., Noida and Verific design Automation, Kolkata from November'2006 to March'2009. Right now, his research areas are Cryptography, Parallel and Distributed Computing.*

Supervisor: Prof. Dipanwita Roy Chowdhury

Algebraic Cryptanalysis of Stream ciphers by Grobner Basis

Cryptanalysis is the art and science of knowing the weaknesses of, if not breaking, the cryptographic ciphers designed. There have been many techniques to do cryptanalysis. Algebraic cryptanalysis is the study of cryptanalytic techniques by employing algebraic means. The basic working principle of algebraic cryptanalysis is to find the equivalent algebraic equations of the cipher under study and then try to solve those equations by algebraic techniques of equation solving. In this work the efforts have been to cryptanalyze the Bivium stream cipher, which is one of the finalists of the eSTREAM project. Grobner basis is a special kind of basis that helps to find out the solution of a system of non-linear equations. In order to handle a large amount of equations a tool has been developed which implements a distributed approach for Grobner basis calculation. A naïve SAT solver has also been written to accomplish the whole task.

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.

We have carried out exhaustive studies over a few students projects that represent few real world problems to evaluate the effectiveness of our approach. From our empirical studies, we observed that our ESRA approach has gained more accuracy in reliability prediction compared to other approaches.

Sourya Bhattacharyya

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

***Sourya Bhattacharyya** received a B.E. degree in Computer Science from Jadavpur University, Kolkata in 2006. From July 2006 till July 2009, he worked in STMicroelectronics, Greater Noida, as a Software Design Engineer in the multimedia domain. 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 Neonatal EEG signal Processing.*

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

Analysis of video-EEG for Newborn Seizure Detection

Human brain cells consist of about 100 billion (10^{11}) neurons. Brain functionalities such as thinking, learning and memorizing are caused by the transfer of electrical impulses through neuron junctions (also called synapses) and interconnections. This neural network in the brain is usually formed within 4 - 5 years of life. Thus, in a mature brain, neuron structure and electrical patterns are more or less consistent; whereas in a newborn brain, these are in the process of evolution. Electroencephalography (EEG) monitors the neural electrical activities in the brain and provides a sensitive real time graphical representation of the brain function. Particularly for neonatal brain, which is vulnerable to changes in oxygenation and blood pressure, EEG can help in diagnosing such physiological changes.

Seizures in the neonates (newborns having birth age at most 4 weeks) are defined as paroxysmal alterations in neurologic functions (such as, behavioral, motor, or autonomic function). This definition of seizure encompasses both clinical phenomena that correlate temporally with epileptiform EEG abnormalities and stereotypical, paroxysmal clinical activities that are not associated clearly with EEG alterations. Seizures are generated due to synchronous neuronal firing in the cerebral cortex. They are associated with underlying conditions such as brain hemorrhage, stroke, meningitis, and hypoxic ischaemic encephalopathy. Neonatal seizures generally have duration less than 2 minutes. However, unlike the seizures in adults or matured children, neonates commonly may not exhibit visible clinical signs and symptoms during seizures.

Neonatal EEG can have long recording duration (even 1 - 2 days). Visual monitoring of the recorded EEG is done by the doctors or technicians for seizure detection. Sometimes, video of the newborn is also recorded and then analyzed to find the correlation between clinical and electrical activities. However, close monitoring of both EEG and video is time consuming and often depends on the subjective assessment of the reviewers. Lack of experts in this field further emphasizes the necessity to develop a tool which can effectively summarize the seizure patterns (or in general the patterns of clinical interest) from the recorded EEG along with the corresponding video segments. With such a tool, doctors can browse through only the marked EEG patterns and summarized video segments for quick diagnosis of the newborn brain. Current research focuses on two problems - 1) Detection of patterns of interest from recorded newborn EEG, and 2) Summarization of the associated video to identify the clinical activities and analyze its correlation with the recorded EEG.

Sandipan Mandal

Email: mandal.sandipan@gmail.com
Joined the department in: July 2009

***Sandipan Mandal** received a 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 2008 to 2009. Since 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

Automatic Speech Recognition in Bengali and Speech-based Application

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. I order to develop a good ASR, we are focusing on optimal text selection technique, grapheme to phoneme conversion, proper phone selection, good quality speech corpus collection and various speech recognition technologies. Initially HMM-based Bengali ASR has been developed using trigram language model and triphone clustering. Gradually we have incorporated syllable level trigram generation and many adaptation techniques in our ASR system. We have also achieved satisfactory result in Bengali phone recognition system and as well as in Bengali syllable recognition from speech input.

The need for voice augmented natural interfaces is clearly important for mobile/embedded devices. However, real-time performance of automatic speech recognition system on embedded platforms is not satisfactory. The reason for such poor performance can largely be attributed to the inherent complexity of speech recognition algorithms, lesser computing power and severe memory constraints to compute the same. So, we are also working on the process of making the ASR fast so that it can be used in embedded platform.

Several optimization techniques including parallelization, input/output operation reduction to reduce cache miss, code level modification, usages of performance optimized compiler flags, beam pruning and other performance tuning to make faster continuous speech recognition. Even syllable level triphone generation in Bengali increases the speed of recognition time.

Voice-based applications are one of the most important area of speech recognition. 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. Handing the voice command, voice search (e.g. Phone Number in SMS) in those application are also very much challenging.

Biswajit Das

Email: biswajit.net@gmail.com
Joined the department in: July 2009

Biswajit Das *received a B.Tech. degree in Electronics and Communication Engineering from Murshidabad College of Engineering & Technology , WBUT, in 2006, and pursuing M.S. degree in Computer Science & Engineering from Indian Institute of Technology, Kharagpur. From January 2007 till January 2008, I worked as a junior Project Assistant in department of Mining, IIT, Kharagpur. Since February 2008, I have been working as Junior Project Officer in the department of Computer Science & Engineering in IIT Kharagpur. My research interests are in the areas of speech processing and speech recognition..*

Supervisor: Prof. Pabitra Mitra

Robust Speech Recognition System for Aged Population

Automatic speech recognition (ASR) system is going to be a useful part of our day to life. ASR system converts speech signal to its corresponding lexicon. ASR system has been developed in many languages, and for different age groups but research on robust ASR system design in Bengali is not progressed yet. My research is directed towards building up an ASR system for aged population. ASR system performs satisfactorily under controlled condition. When, acoustic condition of training and testing differs, ASR performance degrades rapidly. We had developed one baseline ASR system in Bengali with the training data, collected from young population. This system performs well for test data of young people but recognition accuracy decline for test data of aged population. Different effects, which are responsible for mismatch, can be compensated with different normalization or adaptation techniques. I have decided to explore different speaker normalization and adaptation algorithm to improve ASR performance for aged population.

Human articulation system goes through physiological and anatomical changes during their whole life cycle. Those changes have relative effect on spectral properties like formant frequencies, fundamental frequency, jitter, shimmer, harmonic-to-noise ratio and voice-on-set time of speech signal. Perceptual quality of speech signal also degrades with aging.

In my work, I have targeted to find out which acoustic unit (phoneme) is degraded more with respect to speech recognition. In speech recognition, phoneme are wrongly recognized or confused with some other phone. It also happens with the ASR system of young population but confusion increases for aged population. I have done a detailed study on phoneme recognition performance of aged people. It has been observed that speech sample of aged group recognized well with acoustic model of aged people but degrade with acoustic model of young people. Acoustic model developed with training data of young and aged people performs in average. It is not practically possible to collect all variability into training data. To make up these deficiencies, I have used well known vocal tract length normalization (VTLN) and maximum likelihood linear regression (MLLR) techniques. I have used lattice based discriminative acoustic modeling algorithm to increase probability likelihood of the model parameters. I have tried to create a appropriate language model for aged population. It will increase the recognition of day to life conversation.

At the end of my work, I have analyzed the recognition in all aspect. I have verified monophone and triphone acoustic model of aged as well as young user group with test sample of elderly people. To be assured about the divergence between acoustic model of young and aged people, I have applied different statistical distance. Those statistical distances reveal the cause of deterioration of ASR performance.

In future, I will try to adapt the acoustic feature and model according to pronunciation of aged population. I will also try to build a robust language model for all user groups.

Ritwika Ghose

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

Ritwika Ghose *received a 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 Human Computer Interaction and Natural Interface Design.*

Supervisor: Prof. Anupam Basu

Restructuring of a web page to reduce cognitive load of a blind person

Advancements in technology and World Wide Web have made it effortless for people to access wide variety of information of different genres, communicate to people in different parts of the world, perform important transactions like banking or reservation sitting right at home. The Internet caters to the requirements of people from diverse backgrounds and different age levels. With new developments in the field of software and technology, people with various disabilities, like visual handicap, motor disability, etc. have also been capable of exploring the vast facilities of the World Wide Web. However, since most web pages are designed considering the general mass of people without any disability, some accessibility issues may arise when these pages are accessed by a person with some disability. For example, a page with a large amount of content requires scrolling of the page, which is often a difficulty for a person with motor disability. A page containing flashy contents can be a distraction especially for people with cognitive impairment.

This research focuses on improving the web browsing experience for a visually handicapped person. The two modes of accessing text content by a blind person are either by speech or Braille. Accessing documents converted to Braille requires either a Braille tactile device or a Braille printer, which is expensive and difficult to acquire. So the most widely used mode of accessing text content from a computer is via speech. For this purpose, the two-dimensional information presented in a web browser has to be converted to one-dimensional speech. With the advent of Text-to-Speech systems, use of screen readers for reading out information has been popular amongst the blind community. However, since a screen reader reads a page from top to bottom sequentially, reaching a desired portion of the text may take some time. Often a web developer focuses on enhancing the visual appeal of a page, in order to catch maximum attention of the reader. Various types of layout design when presented by a screen reader, often produce confusing screen reader feedback. Poorly designed or unlabeled forms, pictures with no meaningful alt text or inaccessible scripts add to the confusion.

The main aim of this research is to analyze the complexity of a web page, and generate a reading order for the screen reader, so as to reduce the cognitive load for a visually handicapped person. Reading order refers to the direction of the flow of a sequence of text. Reduction of cognitive load in the case of accessing a web page refers to the ease with which a blind user can hear through the contents of a page, with maximum efficiency and minimum distractions. In order to address the above issues, first of all we propose an algorithm to divide a web page into blocks. A block is a logical part of a web page containing a certain type of information unit. For example, a paragraph can be a block, an image with the associated text can be a block, a table can be a block, there can be container blocks which contain other blocks, etc. We define these blocks as a node in a graph structure. Certain attributes of these nodes like position in the page, height, width, alt text (if present), fonts, etc. are noted. Then we define relationships among these nodes, which represent the edges in the graph. These relationships can be physical adjacency, referential links, hyperlinked relationship and contextual adjacency. Physical adjacency refers to neighbor blocks as appearing visually. Referential links refer to the node which is referred to by another node by some particular keyword, like see above, see below, refer to table, refer to figure, etc. Hyperlinked nodes are those which are connected by intra-page hyperlinks. Contextual adjacency refers to those nodes which come in a sequence regarding the same context but may be separated by other nodes which are not related to the context, for example advertisement nodes. The complexity of the page is calculated by giving weights to the nodes and the relationships. For example, the complexity is higher when there are too many different types of nodes clustered in a small area of the page. Complexity is also increased when physical distance between two contextually related nodes is high. For a sighted user, the impact of these factors are less as they can simply skip reading the content in between and directly reach the desired node. But for a blind user, the screen reader would read the content in between, and the user may take some time to realize that this node is irrelevant and then have to jump to the next neighbor node. All these factors lead to the need of rearrangement of the informative content blocks when presented via a screen reader.

Prosenjit Dhole

Email: prosenjit.dhole@gmail.com
Joined the department in: October, 2009

Prosenjit Dhole *received a B.Tech. degree in Computer Science from Institute of Engineering and Management, Salt Lake in 2006. From July 2006 till September 2009, he worked in Interra Systems, Kolkata, as a Senior Member of Technical Stuff. Since January 2010, he has been an MS student in the department of Computer Science & Engineering in IIT Kharagpur. His research interests are in the areas of Delay Tolerant Networks.*

Supervisor: Prof. Arobinda Gupta

Mobility Aided Broadcasting in Delay Tolerant Network

Delay Tolerant Network (DTN) is a type of network primarily characterized by frequent disconnection of end to end paths between nodes. Node mobility is one of the primary reasons for such behavior. Broadcasting is a problem of sending messages from one node to all other nodes in the network. Broadcasting in DTN becomes a challenge as most of the classical broadcasting algorithms fails due the intermittent connectivity of nodes. In real life, node mobility is often repetitive and time sequenced, resulting in predictable contact sequences. We are trying to extract contact patterns from the mobility trace, and exploit those patterns to design an effective broadcast protocol in DTN.

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:

- Dependence of percolation threshold of a superpeer network on its peer degree, superpeer degree at different levels of degree-degree correlation
- Impact of different attacks on the topology of a network at different degree-degree correlation.
- Impact of attacks on the degree-degree correlation of a network.
- Determining the parameters that govern the stability of superpeer networks

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.

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 a 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, Hardware/Software co-verification, Low-power Circuits and System, Power aware High-level Synthesis, Fine grained Power Management strategies.*

Supervisors: Prof. Pallab Dasgupta and Prof. Ajit Pal

Power aware scheduling and binding during high-level synthesis using fine-grained power management strategies

The operator scheduling problem is at the heart of every synthesis tool for digital integrated circuits. Given a data flow graph (DFG), where nodes represent operations and edges represent precedence between operations, the traditional goal of the operator scheduling problem has been to schedule the operations along clock cycles so as to minimize the total number of cycles (representing delay) and the total number of resources (representing area). There exists a huge volume of literature which addresses this optimization problem. In the last decade, power has become a major optimization criterion in all levels of circuit design, ranging from architectural level power management of large power domains to transistor level power management techniques like resizing and adaptive body biasing. While coarse grained power management at the level of architectural power domains (like processor cores and memory banks) has become ubiquitous, the use of fine grained power management at the level of individual functional units is more recent phenomenon. Early work in this area centered around the use of multiple implementations of a given type of functional unit, each of which has the same functionality, but works on a specific voltage (Vdd) and requires a specific number of cycles. These methods demonstrated the trade-off between area and power in the context of operator scheduling.

As static power continues to rise due to the down-scaling trend of CMOS technology, fine-grained power management will have increasing significance in digital circuit design. More recently, fine grained dynamic voltage scaling (FGDVS) technology is used to enable a single functional unit to work with different voltages and in correspondingly different number of cycles. The use of DVS enabled resources reduces the area as compared to Multi-Vdd approaches. Such FGDVS enabled resources implements DVS with local header switches that are inserted down to the sub-block level. The fine-grained header switches add temporal granularity to switch between different processing rates and enables a functional unit to be reused with a different operating voltage, instead of permanant assignment of voltage to components as done in Multi-Vdd enabled functional units. This architecture combines the benefits of Multi- Vdd, fine-grained header switches and dynamic voltage scaling to achieve a wide range of processing rates as well as maximum energy savings. The classical operator scheduling problem during high level synthesis is revisited in the context of FGDVS enabled resources. FGDVS architecture is most beneficial for systems that implement time-wise mutually exclusive functions on the same hardware.

The benefits also increase for DFGs with heterogeneous operations compared to Multi-Vdd where different functions have to be implemented with different components at each voltage level. When the available slack is sufficient enough such that it allows heavy energy consumer operations of the DFG to be scheduled at lower voltages, then fine-grained DVS helps to achieve sufficient gain both in terms of area and power parameters.

The fine-grained power management architecture is suitable for mitigating the increase in energy loss due to leakage. The functional modules that are unused during execution of a given DFG consume leakage energy for single-Vdd and Multi-Vdd approach but FGDVS architecture with component-level header switches could switch-off unused components using power gates leading to minimal leakage during standby or sleep mode. Further, using FGDVS, the total number of resource instances of each type required to schedule a operator graph is often less than the other two approaches, which in turn reduces leakage power as well as makes it more area efficient despite the presence of additional supply rails. However, the main overheads of FGDVS that influence scheduling algorithms on this architecture are Vdd-Switching Power and Vdd-Switching delay. The scheduling decision is made such that the power required to schedule an operation of a DFG at low voltage plus the Vdd-Switching power for switching the unit down and up again must be less than the power consumed for execution of that operation at highest voltage. The algorithm considers these overheads to model power as a cost function to be optimized. The availability of DVS enabled functional units have a significant impact on the size of the state space for the operator scheduling problem.

We explore the extensions of the traditional list based algorithm to handle different area, power budgets and latency constraints. We observe that the algorithm fails to generate a solution in most cases due to these bounds. Moreover it does not show the trade-off between area and power. To address these limitations of list-based algorithms, we present a branch-and-bound formulation for traversing the state space to find solutions lying on the pareto-optimal front with respect to area and power. Since the problem is a multi-objective optimization problem, the pruning criterion is based on dominance, that is, a potential solution is discarded when it has exceeded both the area and power of some previous solution.

We perform our experiment on standard operator scheduling benchmarks. We show that the algorithm is able to report the solutions on the pareto-optimal front within feasible limits of time, which is adequate for practical purposes, since operator scheduling is likely to be done only once in a given design. In this variant of the operator scheduling problem, it is extremely hard to predict the trade-off between the area and power. Therefore systematic traversal of the state space can yield the non-dominated solutions in sequences which may be non-monotonic with respect to both criteria. In such state spaces, constraints on one or more dimensions typically prove to be very useful in pruning the state space. We also explored the performance benefits of using power budget, area budget constraints within the branch-and-bound formulation along with the latency bound.

Binanda Sengupta

Email: binanda.sengupta@cse.iitkgp.ernet.in
Joined the department in: July 2010

Binanda Sengupta *received his B.E. degree in Computer Science and Engineering from Jadavpur University, Kolkata in 2007. From 2007 till 2010, he worked in Tata Consultancy Services Limited, Kolkata, as an Assistant Systems Engineer. Since July 2010, he has been an MS student in the department of Computer Science & Engineering in IIT Kharagpur. His research interest includes Cryptography and Computational Number Theory.*

Supervisor: Prof. Abhijit Das

Parallelization of Different Sieving Techniques

Different sieving techniques are used extensively in solving number-theoretic hard problems like Integer Factorization Problem and Discrete Logarithm Problem. These techniques are incorporated to make the running time sub-exponential (though super-polynomial). We are trying to implement these techniques efficiently and parallelize them in order to reduce the running time.

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 in Indian Institute Technology, Kharagpur, as a Program Facilitator. From July 2008 to October 2009, he has been working as a Junior Project Assistant in the India Chip Design program in IIT Kharagpur. His research interests are in the areas of High-level Synthesis and secure transistor level circuit 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 the behavioral descriptions. To deal with the increasing complexity of today's VLSI designs, the use of HLS tools becomes increasingly crucial. Over the last several years, various such HLS tools have evolved producing elementary non-optimized data paths to more sophisticated one generating data paths optimized with 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 give us 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 forced the SAST to execute the input behavior on that architecture. 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 and design space can be explored by simply changing the architectural parameters for the same input behaviour. 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 industrial tools like 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.

Presently, SAST does not support global memory. As a result, SAST may not generate optimize design for array intensive applications. My next objective is to extend SAST implementation 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 objective is to evolve the suitable architecture and subsequently tune the phases of HLS process accordingly for array intensive programs.

Secure multiplier design to counter side channel attack (differential power analysis)

Side channel attack is any attack based on the information gained from the physical implementation of the cryptographic system rather that by brute force or theoretical weakness in cryptographic algorithms. Timing information, power consumption electromagnetic leaks, fault injection or even sound can provide extra source of information which can be exploited to break the system. Differential power analysis is one of the method of side channel attack.

Differential power analysis: It involves in collecting many power traces and performing statistical analysis of the power variation with respect to changes in data values and poses a serious threat so the cryptographic devices.

My aim is to build three side channel (Differential power) resistant asynchronous multiplier using three different multiplication algorithms namely Booth's multiplication, Wallace tree based multiplication and Karatsuba based multiplication. My objective is also examine the power after Fabricating those multiplier using UMC 180nm technology.

Suman Kalyan Maity

Email: sumankalyan.maity@cse.iitkgp.ernet.in
Joined the department in: July 2011

***Suman Kalyan Maity** has received a B.Tech. degree in Computer Science & Engineering from National Institute of Technology, Durgapur in 2011. Since July 2011, he has been a research scholar in the department of Computer Science & Engineering in IIT Kharagpur. His research interests are in the areas of Complex Systems and Language Dynamics.*

Supervisor: Prof. Animesh Mukherjee

Opinion formation in time-varying social network: The case of Naming Game

Social networks are inherently dynamic. Social interactions and human activities are intermittent, the neighborhood of individuals moving over a geographic space evolves over time, links appear and disappear in the World-Wide-Web. The essence of social network lies in its time-varying nature. Links may exist for a certain time period and may be recurrent. In summary, as time progresses, the societal structure keeps changing. Similarly, with the evolution of time, social conventions, shared cultural and linguistic patterns reshape themselves. Opinions spread, some gets trapped into communities, some crosses the barrier of local groups/communities and become accepted globally among different communities and some die competing with others. Most of these social phenomena can be modeled and analyzed in a time-varying framework.

We focus on the basic Naming Game model (NG) (Baronchelli et al., 2006b) to study how opinions spread with time and how societies move towards consensus in the adoption of a single opinion through negotiation or agree upon multiple opinions due to non-uniform interaction pattern among different communities. The evolution of the system in this model takes place through the usual local pairwise interactions among artificial agents that necessarily captures the generic and essential features of an agreement process. This model was expressly conceived to explore the role of self-organization in the evolution of languages and has acquired a paradigmatic role in semiotic dynamics that studies evolution of languages through invention of new words, grammatical constructions and more specifically, through adoption of new meaning for different words. The NG finds wide applications in various fields ranging from artificial sensor network as a leader election model to the social media as an opinion formation model.

The basic NG is played by a population of agents in pairwise interactions to negotiate conventions. At each time step, a randomly chosen speaker voices a random opinion from his list to a randomly chosen neighbor, designated as the hearer. If the hearer has the spoken opinion in his list, both the speaker and the hearer retain only that opinion removing all other competing opinions from their respective inventories; else the hearer adds the spoken opinion to his list.

We have studied the NG in time-varying scenarios of 69 days of face-to-face interaction data from Science Gallery visitors (http://www.sociopatterns.org/datasets/) where we choose the speaker randomly and preferentially select the hearer from its neighbors. The interesting observation is the time to reach the global consensus (the state when every agents agrees upon a single word) . The consensus time is very long for those instances of daily network. The reason behind this is the inherent community structure of social networks. Each community reaches internal consensus fast but the weak connections between communities are not sufficient for opinions to propagate from one community to the other leading to long multi-opinion states which are also known as "metastable states" in the domain of statistical physics . Presence of community structures slows down the dynamics, however, what renders the system even slower is the presence of different-sized communities. The reason for this is quite straight-forward: the agents that are part of a larger size community have a higher probability of being chosen for a game than those belonging to a smaller size community. This is a reminiscent of the fact that the agents are chosen randomly which automatically increases the chances of landing in a larger size community simply because a larger bulk of the population is confined within this community. Therefore, even when consensus is reached very fast in a large community, the system keeps on choosing agents from this community itself mostly resulting in "success with no outcome".

Further, we investigate the Naming Game dynamics in perfect synchronization with an evolving social Network (Hypertext Conference Dataset ,2009) shedding new light on the basic emergent properties of the game that differs largely from what is reported in the existing literature.

References

[1] A. Baronchelli, M. Felici, V. Loreto, E. Caglioti, and L.Steels, 2006b, Sharp transition towards shared vocabularies in multi-agent systems, Journal of Statistical Mechanics: Theory and Experiment, Vol. 6, No. 6, 2006

Indrasish Saha

Email: indrasish88@gmail.com
Joined the department in: July 2011

Indrasish Saha *received a B.E. degree in Information Technology from Jadavpur University, Kolkata in 2011. Since July 2011, he has been a research scholar in the department of Computer Science & Engineering in IIT Kharagpur. His research interests are in the areas of Hardware Security.*

Supervisor: Prof.Rajat Subhra Chakraborty and Prof.Debdeep Mukhopadhyay

Model Building Attacks on Physically Uncloneable Functions

A Physical Random Function or Physical Unclonable Function (PUF) is a function that is based on a physical system that is easy to evaluate (using the physical system) and its output looks like a random function which is unpredictable even for an attacker having physical access. In the context of intrinsic physical properties of integrated circuits, Physically Unclonable Functions (PUFs) can be used to complement classical cryptographic constructions, and to enhance the security of cryptographic devices. PUFs have recently been proposed for various applications, including anticounterfeiting schemes, key generation algorithms, and in the design of block ciphers.

Numerical modeling attacks on PUFs presume that an adversary Eve has collected a subset of all Challenge-Response pairs of the PUF, and tries to derive a numerical model from this data, i.e. a computer algorithm which correctly predicts the PUF's responses to arbitrary challenges with high probability. If successful, this breaks the security of the PUF and of any protocols built on it. We have collected Challenge-Response pairs for Arbiter and Ring Oscillator based PUFs and propose to build machine learning techniques for such modeling attacks.

Tamal Sen

Email: tml.cse@gmail.com

Joined the department in: December, 2011

***Tamal Sen** received a BTech (2008) degree in Computer Science and Engineering from Jalpaiguri Govt. engg. College. From Feb'09 to Nov'11 he worked in Cognizant Tech. Solutions, Kolkata. Since Dec'11, he has been an MS scholar in the department of Computer Science & Engineering in IIT Kharagpur. His research interests are in the areas of program analysis and effort minimization in software testing.*

Supervisor: Prof. Rajib Mall

Regression Test Selection for Component-Based Software

A component is a cohesive group of reusable services that have been collated into an executable unit. Components are developed independently, available off-the-shelf and integrated into a component based system by the application developers. Component-based development has found rapid acceptance by software developers since it is realized that reuse of components helps lower the development cost and speed up the development process.

Components can be written in different programming languages and may be distributed across different platforms. Source code is not usually disclosed to the application developers. Components can have two different types of operations: provided and required. Provided operations are those which have been implemented in the component and they can be invoked from the outside. In the other hand, concrete implementation of the required operations is left to the users of the component. The provided and required operations are published through the interfaces called required interfaces and provided interfaces respectively. Interfaces are generally defined using an interface definition language (IDL). The application developers (the users of the components) make use of interface definitions of the components to develop an application.

A major difficulty in component based development is that unavailability of source code makes software engineering tasks difficult to carry out. In addition to the development activities, testing activities like coverage analysis, regression test selection etc. turn out to be even more challenging.

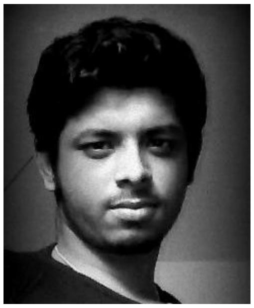
Regression testing is carried out after every code modification to ensure that the unmodified functionalities continue to work satisfactorily. Regression test selection (RTS) is the process of selecting a subset of system test cases which can effectively find all errors that might get induced in the unmodified parts.

In case of component-based systems, new versions of the components are released very frequently. In this context, analyzing change impact and selecting a safe subset of the system test cases for regression testing involve high overhead, thereby pose significant challenge.

For component-based software, applying traditional RTS techniques is difficult because application developers do not have the source code for analyzing change impact; neither can they obtain coverage data of the test suite through code instrumentation. It may be unrealistic to expect component vendors to provide that information due to obvious reasons. To facilitate regression testing, component vendors either need to provide Built-In-Test interfaces or after each modification, they are expected to provide the change information in terms of some published elements (such as method signature, pre/post conditions etc.).

The simplest form of change information can be a collection of affected methods which could be published after every modification to a component. Techniques for identifying affected methods can be simply choosing those methods which have been modified or those that directly or indirectly call a modified method. But errors can exist in the unmodified parts of the code due to its dependencies on the actually modified parts. It is suggested that analysis of control and data dependence relationships is necessary for detecting many types of code-based faults. Hence, the methods which may execute some of those indirectly affected statements need to be included in the change information in order to perform a better regression testing.

In contrast to what is implicitly assumed by many existing techniques, invoking an affected method does not ensure that all affected statements inside that method will be executed. It can be argued that, since most components have non-trivial state models, some statements inside a method may remain unreachable in some states, even if the method is invoked in all possible ways. Pure dependence based techniques choose test cases which invoke one or more component methods which have been found affected by dependence analysis. Some test cases, chosen by pure dependence based technique, can be redundant if they invoke affected methods in a state such that no affected statements are reachable. We propose an RTS technique for component based software which includes an enhanced mode of publishing change information by statically analyzing the source code of a component. The approach can reduce regression testing effort significantly by reducing the number of regression cases.



Parnab Kumar Chanda

Email: parnab.2007@gmail.com
Joined the department in: January 2012

***Parnab Kumar Chanda** received a B.E. degree in Information Technology from School Of Information Technology(Formerly IIIT-Kolkata) ,WBUT Kolkata in 2009, From Jan 2010 till July 2011, he worked in Infosys Technologies, Chennai, as a Software Engineer. Since January 2012, he has been aresearch scholar in the department of Computer Science & Engineering in IIT Kharagpur. His research interests are in the area of Information Retrieval.*

Supervisor: Prof. Sudeshna Sarkar

Cross Language Information Access

Web documents are growing with multilingual content every day. Since different medias, in countries like India and Europe, share the information like news, blogs, cinema, etc in the regional languages of the people, it becomes essential for the users to access the information rich content present across languages. Thus the need for a Cross Lingual Information Access (CLIA) system grows faster. Such systems would be expected 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. At present, we are looking at such a cross lingual information access system that is specifically intended for retrieval of documents written in Indian languages pertaining to the tourism domain. The basic idea of the CLIA system is as follows: A user is expected to choose any one of the 9 selected Indian languages or English and fire a query written in the language of their choice seeking certain information related to the tourism domain. CLIA system retrieves top k documents pertaining to the given user query and presents the ranked list of documents sorted by their similarity scores.

The overall structure of the system can be seen into two parts: offline components and online components.

Offline components include: a web crawler (fetches html documents from the world wide web), parser (removes the noisy content and extracted the filtered text content with meta tags), language identifier (identifies the language of the document), domain specific document classifier (a document classifier to check whether the given document belongs to tourism or health or others), language specific stemmers and stop word remover (both for specific Indian languages), and indexer (to create an inverted index of the parsed web documents, each having specified fields).

Online components include: GUI (to feed the user queries by choosing the language of the choice of users), query processing module (forms the expanded query from the user keywords with specified boost factors for the fields), searcher (uses the ranking strategy that computes the similarity between the query and documents), output presentation (snippet generation and results page generation with further navigational links).

The ranking of web documents will be the primary focus in the current work. At present, the focus is on identifying the underlying topic(s) of the extracted content of each web document and then modeling and applying the ranking function using this topic information as a feature with various similarity measures. The final ordering of documents would be based on the similarity scores computed by the defined ranking function that maps the document topics with the actual intent of the user query.

Ayan Palchaudhuri

Email: ayan@cse.iitkgp.ernet.in
ayanpalchaudhuri@gmail.com
Joined the department in: February 2011

Ayan Palchaudhuri *received a B.Tech. degree in Electronics and Communication Engineering from West Bengal University of Technology in 2010. From February 2011, he worked as a Junior Project Assistant in the Department of Computer Science & Engineering, IIT Kharagpur, under the project : Hardware Security : Ensuring Trust in Integrated Circuits. Since December 2011, he has been a research scholar in the department of Computer Science & Engineering in IIT Kharagpur. His research interests are in the areas of VLSI and Reversible Watermarking of Digital Images.*

Supervisor: Prof. Rajat Subhra Chakraborty

FPGA based Hardware Design for Real-time Reversible Watermarking of Digital Images

Digital watermarking has been widely used to protect the copyright of digital images. In order to strengthen the intellectual right of a digital image, a trademark of the owner could be selected as a watermark and embedded into the protected image. Then the watermarked image could be published and the owner can prove the ownership of a suspected image by retrieving the watermark from the watermarked image. For some critical applications such as the law enforcement , medical and military image system , it is crucial to restore the original image without any distortions. The watermarking techniques satisfying these requirements are referred to as ‘reversible watermarking’. Reversible watermarking is designed so that it can be removed to completely restore the original image.

Watermarking implementations can be done in software or in hardware. Although it might be faster to implement an algorithm in software, there are a few compelling reasons for a move towards hardware implementation. A hardware watermarking solution is often more economical because adding the watermarking component takes up a small dedicated area of silicon. In software, implementation requires the addition of a dedicated processor such as a DSP core that occupies considerably more area, consumes significantly more power, and may still not perform adequately fast.

Hardware implementations of watermarking can be implemented in Application Specific Integrated Circuits (ASICs) or in Field Programmable Gate Arrays (FPGAs). Most of the current hardware implementations have been done for ASIC designs. Recent advances in FPGA technology, such as 90nm process devices, higher gate densities, better interconnect architectures, reduction in power consumption, multiple I/O formats and embedded optimized logic, have allowed for applications that were previously intended for ASICs to be implemented in FPGA devices, with the added value of a lower FPGA cost when compared to an ASIC.

Arnab Dhar

Email: arnab832007@gmail.com
Joined the department in: January 2012

***Arnab Dhar** received a B.Sc. degree in Computer Science from Asutosh College, Kolkata in 2004, and an MCA degree from RCC Institute of Information Technology, Kolkata in 2008. He worked in CVPRU, ISI, Kolkata, as a Project Linked Person for 2 years. Since July 2011, he has been working as a Junior Project Officer in ILMT project in IIT Kharagpur. He has joined in MS (by research) programme of Computer Science & Engineering Department, IIT Kharagpur, in January 2012. His research interests are in the areas of Computational Natural Language Processing.*

Supervisor: Prof. Sudeshna Sarkar

Bangla Dependency Parser

Dependency parsing is the automatic analysis of the dependency relations in natural language sentences. The nodes of the parse tree represent the words and edges represent the dependency relations. There is a one to one correspondence between the parse tree and the sentence. Dependency relations are defined as the binary syntactic-semantic relations between the words. In recent years, Indian language dependency parsing gained a lot of attention and popularity. The parsers are used in almost all natural language applications like Machine translation, Summarization, Information retrieval, etc.

Dependency parsing can be broadly divided into grammar-driven and data-driven parsing. Many of the modern grammar-driven dependency parsers parse by satisfying the given set of constraints. Data-driven parsers, on the other hand, use a dependency tagged corpus (Treebank) to induce a probabilistic model for disambiguation. There are several statistical parsers available that can automatically create the model from the Treebank. However, some Bangla linguistic features like, Root, POS category, gender, number, person etc. should be used to observe their performance on Bangla parsing.

Building the dependency parser for a language like, Bangla is challenging due to its morphological richness and relatively free word order properties. The resource required for the Bangla dependency parsing is small as compared to English and European languages. Preparation of the comprehensive dependency relation set and the large sized Treebank for Bangla is still under construction.

Souvik Kolay
Email: souvik1809@gmail.com
Joined the department in: July, 2011

***Souvik Kolay** received a B.Tech. degree in Information Technology from RCC Institute of Information Technology, Kolkata in 2011. Since July 2011, 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.*

Supervisor: Dr. Debdeep Mukhopadhyay.

Design and Analysis of a Light Weight Cryptographic System on FPGAs

Pervasive computing or ubiquitous computing is the growing trend towards embedding microprocessors in everyday objects so that they can communicate information. Pervasive devices, like RFID, possess very limited resources in terms of memory, computing power and battery supply, but many applications running on these devices will contain sensitive information and there lies the need of cryptographic systems to ensure its security. This area of cryptography which deals with the design, analysis and implementation of cryptographic algorithms for devices with extremely constrained resources is formally termed as lightweight cryptography. The current objective of the project is to Design a Lightweight Block Cipher, which will be light enough for implementing in FPGA and at the same time will provide the desirable security. A new lightweight block cipher, named “Khudra” with key size of 80 bit , block size of 64 bit and with the approximate LUT count of 250, is proposed, which can be considered as one of the most compact design among the existing lightweight block cipher which also posses desirable security for the lightweight device.

Debapriya Basu Roy

Email: dbroy24@gmail.com
Joined the department in: July 2011

Debapriya Basu Roy *received a B.Tech. degree in Electronics & Communication Engineering from RCC Institute of Information Technology, Kolkata in 2011. Since December 2011, 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 VLSI design.*

Supervisor: Dr. Debdeep Mukhopadhyay.

Design of Secured Processor for Finite Field

The present project is aimed at designing a finite field controller. The controller is to be optimally designed to perform operations in Galois Fields on FPGAs. Given the large applications of Galois fields in cryptography and communications the design is supposed to be suitably optimized for high performance. Further the requirements of crypto-processors poses more challenges for secured design methodologies.

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*

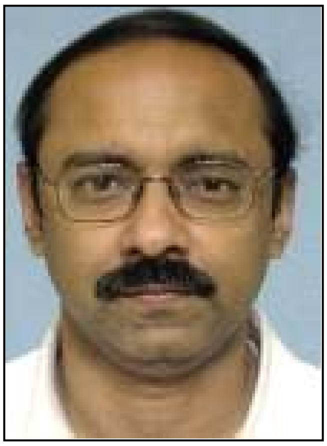
Jayanta Mukhopadhyay received his B.Tech, M.Tech, and Ph.D. degrees in Electronics and Electrical Communication Engineering from the Indian Institute of Technology (IIT), Kharagpur in 1985, 1987, and 1990, respectively. He joined the faculty of the Department of Electronics and Electrical Communication 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 160 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. degree 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

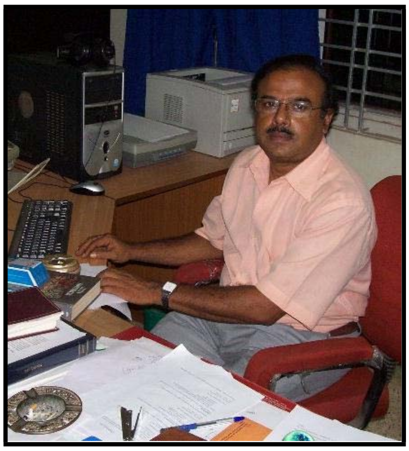


Arobinda Gupta

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 199 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, and 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, and 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, and 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 Defense 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 and 'Formal methods for power intent verification', sponsored by Synopsis (India) Pvt. Ltd. He has over 135 publications in reputed journals and conference proceedings and two books entitled `Microprocessors: Principles and Applications' published by TMH (1990) and `Microcontrollers: Principles and Applications' published by PHI (2011). 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.

Animesh Mukherjee

Email: animeshm@cse.iitkgp.ernet.in

Research Interests: *Complex systems, language dynamics, social computation, web social media.*

Animesh Mukherjee is an assistant professor at the Department of Computer Science and Engineering, IIT Kharagpur. He completed his doctoral degree from the same department and then moved to ISI Foundation, Torino, Italy as a post doctoral researcher. His areas of interest center around studying various complex sociocultural phenomena (e.g., language) under the lens of statistical physics as well computer science. Dr. Mukherjee received the prestigious Young Scientist award from the Indian Science Congress Association in the year 2006. He has authored more than 40 research articles (journals and refereed conferences) and co-edited a book on "Dynamics on and of Complex Networks: Applications to Biology, Computer Science, Economics, and the Social Sciences" to be published by Birkhauser, Springer and a special issue on "Network Models for Cognitive and Social Dynamics of Language" from Computer Speech and Language Journal. He also actively serves in program committees of various top-tier conferences, frequently delivers talks and tutorials and arranges national/international workshops to popularize his field of research.

Chittaranjan Mandal

Email: chitta@cse.iitkgp.ernet.in

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

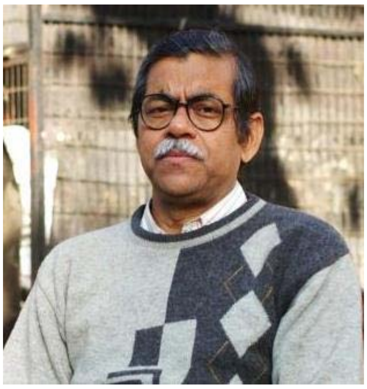
Chittaranjan Mandal received his Ph.D. degree from IIT, Kharagpur, India, in 1997. He is currently a Professor with the Department of Computer Science and Engineering and also the School of Information Technology, IIT, Kharagpur. Earlier he served as a Reader with Jadavpur University. His research interests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

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 Science and Engineering Department from June 2009. Prior to this he worked as an Assistant Professor in the Dept of Computer Science and Engineering, IIT Madras. Debdeep obtained his B.Tech from the Dept of Electrical Engineering, IIT Kharagpur in 2001. Subsequently he obtained his MS Degree in 2004 and PhD from the Dept of Computer Science and Engineering, IIT Kharagpur in 2007. He has authored about 20 Journal and 50 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) Techno Inventor award for best PhD Thesis in 2008, Indian National Science Academy (INSA) Young Scientist Award in 2010, Indian National Academy of Engineers (INAE) Young Engineer Award 2010, and Associates of Indian Academy of Science in 2011. Debdeep has been the recipient of Outstanding Young Faculty Fellowship in 2011 from IIT Kharagpur.



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 Engineering and PhD in Engineering from I.I.T., Kharagpur. He has served I.I.T., Kharagpur as a faculty member since 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 Science and Engineering, Indian Institute of Technology, Kharagpur, India. She received her B.Tech and M.Tech. degrees in Computer Science from University of Kolkata in 1987 and 1989 respectively, and the PhD degree from the department of Computer Science & Engineering, Indian Institute of Technology, Kharagpur, India in 1994. Her current research interests are in the field of Cryptography, Error Correcting Code, Cellular automata and VLSI Design & Testing. She has published more than 140 technical papers in International Journals and Conferences. Dr. Roy Chowdhury has supervised 11 PhD and 8 MS thesis and she is the Principal Investigator of several R&D projects. She is the recipient of INSA Young Scientist Award and Associate of Indian Academy of Science. 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 26 years of teaching and research experience, and over 100 publications in international journals and conferences, and has 13 PhD guidance to his credit. 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, online social networks(OSN), delay tolerant network etc. He has worked on various aspects of OSN like understanding the importance of link farming in OSN and how to discover experts in OSN. 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. 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 100 papers in international conferences and journals. He has also edited a book on Complex Networks published by Birkhauser, Boston. He currently publishes in various top ranking international journals and conferences including ACM CCS, PODC, SIGCOMM, ACL, WWW, INFOCOM, Euro Physics Letters, Pysical Review E, ACM and IEEE Transactions, etc. For more information, please visit http://www.facweb.iitkgp.ernet.in/~niloy

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.*

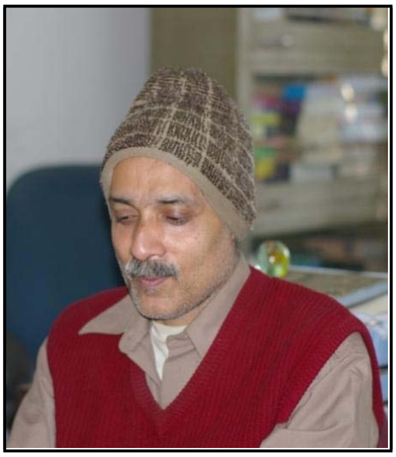
Dr. Pallab Dasgupta did his B.Tech, M.Tech and PhD in Computer Science from the Indian Institute of Technology Kharagpur. He is currently a Professor at the Dept. of Computer Sc. & Engg, I.I.T. Kharagpur. His research interests include Formal Verification, Artificial Intelligence and VLSI. He has over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

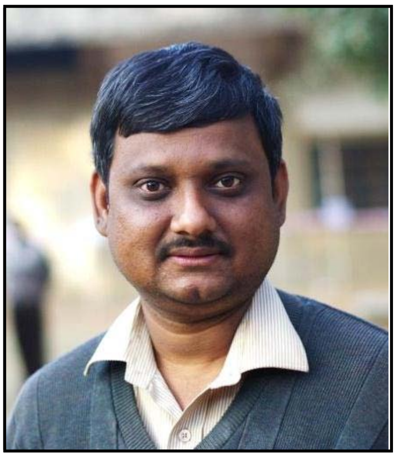


Partha Sarathi Dey

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.

Partha Pratim Das

Email: ppd@cse.iitkgp.ernet.in

Research Interests: Image Processing, Software Engineering, and Embedded Systems.

Dr. Partha Pratim Das obtained his B Tech, M Tech and Ph D degrees in 1984, 1985 and 1988 respectively from the Department of Electronic and Electrical Communication, IIT Kharagpur. He served as a faculty in the Department from 1988 to 1998 and guided 5 PhDs. In 1998, he moved to the Industry and initially worked as a Business Development Manager for Alumnus Software Ltd for over two years before joining Interra Systems, Inc as a Senior Director and Center Head, Kolkata in 2001. In 2011, Dr. Das joined back the Department as a Professor. Since 2003, Dr. Das also serves as a Visiting Professor with Institute of Radio Physics & Electronics, Calcutta University.

Dr. Das has received several recognitions including UNESCO/ROSTSCA Young Scientist Award (1989), INSA Young Scientist Award (1990), Young Associate-ship of Indian Academy of Sciences (1992), UGC Young Teachers’ Career Award (1993), INAE Young Engineer Award (1996), Interra 5 Years’ Tenure Plaque (2007) and Interra Special (Process) Recognition (2009). He is an active member of the VLSI community and has served as General Chair for International Conference on VLSI Design & Embedded Systems, 2005 and as Organizing Chair for International Symposium on VLSI Design & Test, 2007. He also works as a Review Writer for ACM Computing Surveys and is a reviewer for Pattern Recognition Letters. Dr. Das is a member of Association of Computing Machinery and VLSI Society of India.

Dr. Das has published over 40 technical papers in international journals in areas of Digital Geometry, Image Processing, Parallel Computing and Knowledge-based Systems.

Rajat Subhra Chakraborty

Email: pabitra@cse.iitkgp.ernet.in

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

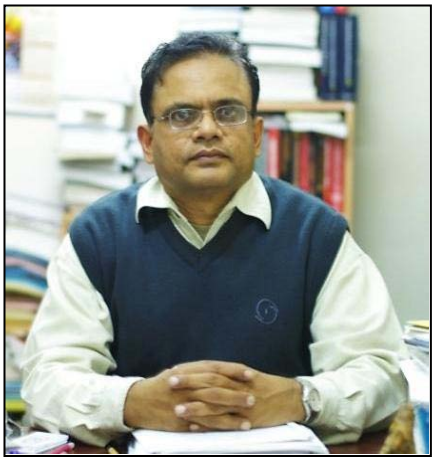
Rajat Subhra Chakraborty is an Assistant Professor in the Computer Science and Engineering Department of IIT Kharagpur. He received his Ph.D. degree in Computer Engineering from Case Western Reserve University (Cleveland, Ohio, USA) in 2010 and a B.E. (Hons.) degree in Electronics and Telecommunication 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. As a graduate student, he has received multiple student awards from IEEE and ACM, and an annual award for academic excellence from Case Western Reserve University in 2009. Part of his Ph.D. research work has been the subject of a U.S. patent filed by Case Western Reserve University in 2009. 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 reversible watermarking for digital content protection. He has over 30 publications in international journals and conferences of repute, (including IEEE TCAD, IEEE TCAS-I, ACM TETCS, IET CDT, IET IP, ICCAD, DATE, CHES, VTS, VLSID, ISQED, HOST, ICISS, ATS, etc.), and has presented his research work at many of these conferences. He has delivered tutorials on Hardware Security at several conferences and workshops, such as: IEEE VLSID, Chennai, India, 2011; IEEE WIFS, Foz do Iguacu, Brazil, 2012; IEEE ATS, New Delhi, India, 2012; ICISS, Kolkata, India, 2012, etc. He has acted as a reviewer and acted as a program committee member for multiple international conferences and journals. He is the co-author of one book on hardware security (forthcoming). Dr. Chakraborty is a member of IEEE.

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.

Research Scholars Graduated in 2011-2012

PHD SCHOLARS

Name: Bivas Mitra
Thesis Title: Analyzing the Resilence and Emergence of Superpeer Networks.
Supervisors: Prof. Sujoy Ghose and Prof. Niloy Ganguly

Name: Najumudheen ESF
Thesis Title: Test Coverage Analysis of Object- Oriented Programs.
Supervisor: Prof. Rajib Mall

Name: Santosh Ghosh
Thesis Title: Design and Analysis of Pairing Based Cryptographic Hardware for Prime Fields.
Supervisors: Dr. Debdeep Mukhopadhyay and Prof. Dipanwita Roychowdhury

Name: Plaban Kumar Bhaumik
Thesis Title: Studies on Multi-Label Classification of Emotions Evoked from Text Data.
Supervisors: Dr. Pabitra Mitra and Prof. Anupam Basu

Name: Gopal Paul
Thesis Title: Studies in Logic Synthesis, Testing, and Security Issues of Digital Circuits Based on Binary Decision Diagrams.
Supervisors: Prof. Ajit Pal and Prof. Bhargab B. Bhattacharyay

Name: Rajarshi Pal
Thesis Title: Modeling and Application of Visual Saliency.
Supervisors: Prof. Jayanta Mukhopadhyay and Dr. Pabitra Mitra

Name: Subrata Nandi
Thesis Title: Information Management in Large Scale Networks.
Supervisors: Prof. Ajit Pal and Prof. Niloy Ganguly

Name: Sourav Das
Thesis Title: Design and Analysis of Stream Ciphers Using Cellular Automata.
Supervisor: Prof. Dipanwita Roychowdhury

Name: Soumyajit Dey
Thesis Title: Formal Analysis of Heterogeneous Embedded Systems Using Tagged Signal Models
Supervisors: Prof. Anupam Basu and Prof. Dipankar Sarkar

Name: Dixit Manoj Gangadhar
Thesis Title: Formal Methods for Early Time-Budgeting in Component Based Embedded Control Systems.
Supervisors: Prof. Pallab Dasgupta and Dr. S. Ramesh

MS SCHOLARS

Name: Debabrata Dey
Thesis Title: Cellular Automata Based Non Linear Stream Cipher.
Supervisor: Prof. Dipanwita Roychowdhury

Name: Debkumar Patra
Thesis Title: Reliability and Interoperability in Distributed Telemedicine Systems.
Supervisors: Prof. Arun Kumar Majumder and Prof. Jayanta Mukhopadhyay

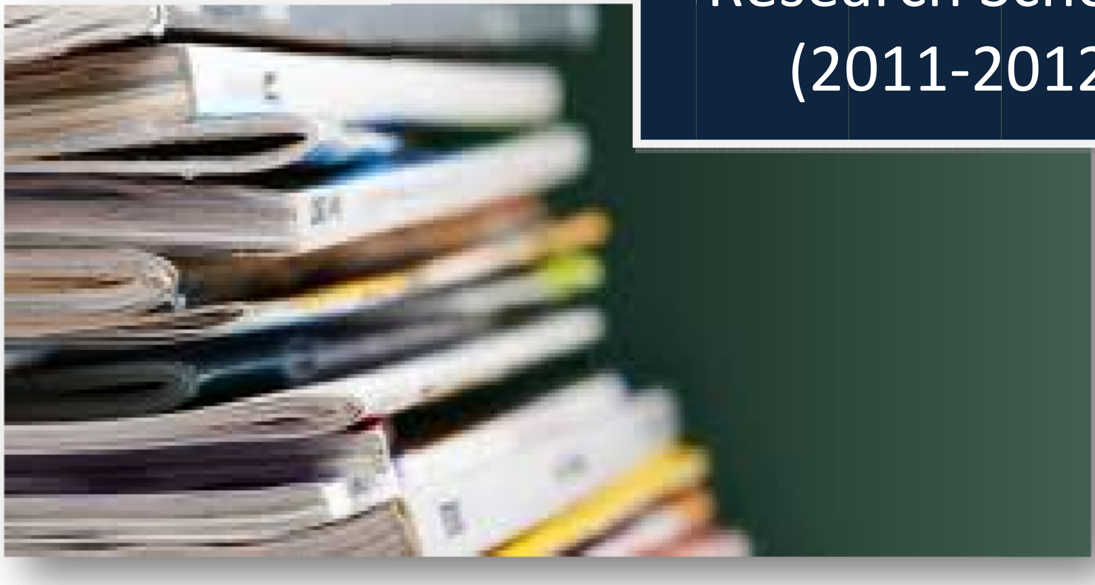
Name: Mayur Bubna
Thesis Title: Some Problems in Testing of Reversible Circuits.
Supervisor: Prof. Indranil Sengupta

Name: Souvik Bhattacharjee
Thesis Title: Parallel and Distributed Implementations of the Lancozs Spares System Solver over Large Prime Fields.
Supervisor: Dr. Abhijit Das

Name: Swarnendu Biswas
Thesis Title: Model-Based Regression Test Selection and Optimization for Embedded Programs.
Supervisor: Prof. Rajib Mall

Name: Anup Kumar Bhattacharya
Thesis Title: Efficient Software Implementation of Elliptic-Curve Pairing
Supervisors: Prof. Dipanwita Roychowdhury and Dr. Abhijit Das

Publications by Research Scholars (2011-2012)



2012

1. A. Hazra, P. Dasgupta and P. P. Chakrabarti; "*Cohesive Coverage Management Leveraging Formal Test Plans*"; LAP LAMBERT Academic Publishers, January 2012, (ISBN: 978-3-8473-7645-3).

2. A. Hazra, P. Dasgupta, A. Banerjee and K. Harer; Formal Methods for Coverage Analysis of Architectural Power States in Power-Managed Designs; In *the 17th Asia and South Pacific Design Automation Conference (ASP-DAC)*, pp. 585-590, January 2012.

3. A. Hazra, S. Goyal, P. Dasgupta and A. Pal, "Formal Verification of Architectural Power Intent", *IEEE Transactions on Very Large Scale Integration Systems* (TVLSI), 2012, (Accepted).

4. C. Karfa, D. Sarkar, C Mandal. "Formal Verification of Code Motion Techniques using Dataflow Driven Equivalence Checking", in *ACM Transactions on Design Automation of Electronic Systems*, 2012,(Accepted).

5. D. Dash, A. Bishnu, A. Gupta, and S. C. Nandy, "Approximation Algorithms for Deployment of Sensors for Line Segment Coverage in Wireless Sensor Networks", In *International Conference on Communication Systems and Networks*, Banglore, India, 2012.

6. D. Dash, A. Bishnu, A. Gupta, and S. C. Nandy, "Finding the Quality of Line Coverage of a Sensor Network", In *International Conference on Distributed Computing and Networking*, Hong Kong, China, January, 2012.

7. D. P. Dogra, A. K. Majumdar, S. Sural, "Evaluation of Segmentation Techniques Using Region Area and Boundary Matching Information". *Journal of Visual Communication and Image Representation (JVCIR)*, Elsevier, Vol. 23, No.1,pp. 150-160, January, 2012.

8. D. P. Dogra, A. K. Majumdar, S. Sural, J. Mukherjee, S. Mukherjee, A. Singh, "Toward Automating Hammersmith Pulled-To-Sit Examination of Infants using Feature Point based Video Object Tracking". *IEEE Transactions on Neural Systems and Rehabilitation Engineering (TNSRE)*, Vol. 20, No. 1, pp. 38-47, January, 2012.

9. D. Pal, P. Dasgupta and S. Mukhopadhyay, A Library for Passive Online Verification of Analog and Mixed-Signal Circuits, In the proceedings of IEEE VLSI Design Conference, pp. 364-369, January 2012.

10. D. Sinha, A. Basu, "Design and Evaluation of a Cognition Aware File Browser for Users in Rural India ", In *Proceedings of Perception and Machine Intelligence 2012*, Lecture Notes in Computer Science, Volume 7143/2012, pp. 129-136, 2012.

11. Krishna Kumar S., S. Kundu, and S. Chattopadhyay, “Customizing completely specified pattern set targeting dynamic and leakage power reduction during testing”, *VLSI Journal*, INTEGRATION, 2012, (Accepted).

12. M. Shirole, K. Mounika, and R. Kumar, “Transition Sequence Exploration of UML Activity Diagram using Evolutionary Algorithm", In *Proceedings of the 5th India Software Engineering Conference, ISEC '12*, 2012, (Accepted).

13. P. Ghosh, A. Hazra, R. Gonnabhaktula, N. Bhilegaonkar, P. Dasgupta, C. Mandal and K. Paul, “POWER-SIM : An SOC Simulator for Estimating Power Profiles of Mobile Workloads”, *Journal of Low-Power Electronics* (JOLPE), 2012, (Accepted).

14. R. R. Maiti, N, Ganguly, A. Gupta, “Epidemic Broadcasting in DTNs using Directional Antenna”, in Posters of the *Fourth International Conference on Communication Systems and Networks (COMSNET)*, 2012.

15. R. R. Maiti, S. Gandhi, N. Ganguly, “Towards Modeling Realistic Human Mobility”, In *the Fourth International Conference on Communication Systems and Networks (COMSNET), PhD Forum*, Bangalore, India, 2012.

16. R. Saha Roy, M. Choudhury and K. Bali, “Are Web Search Queries an Evolving Protolanguage?”, In *Proceedings of the 9th International Conference on the Evolution of Language* (Evolang IX), Kyoto, Japan, pp. 13-16, March 2012, (Accepted).

17. S. Bag, P. Bhowmick, and G. Harit, “Detection of structural concavities in character images–A writer-independent approach”, In *Indo-Japan Conference on Perception and Machine Intelligence (PerMIn)*, Kolkata, India, 2012.

18. S. Bhattacharyya, A. Biswas, J. Mukherjee, A. Majumdar, B. Majumdar, S. Mukherjee, and A. Singh, "*Feature Selection for Automatic Burst Detection in Neonatal Electroencephalogram*", In IEEE Journal on Emerging Topics in Circuits And Systems, Volume 1, Issue 4, pp. 469 – 479, 2011.

19. S. Dhal, I. Sengupta, "A New Authentication Protocol for Multi-tag RFID", In *International Conference on Recent Advances on Information Technology (RAIT) 2012*, ISM Dhanbad, 2012.

20. S. Ghosh, B. Viswanath, F. Kooti, N. Sharma, G. Korlam, F. Benevenuto, N. Ganguly, K. Gummadi, “Understanding and Combating Link Farming in the Twitter Social Network”, In *ACM World Wide Web Conference (WWW)*, Lyon, France, April 2012, (Accepted).

21. S. K. Dandapat, B. Mitra, R. Roychoudhury, and N. Ganguly, "Smart Association Control In Wireless Mobile Environment Using Max-Flow", *IEEE Transaction on Network and Service Management (TNSM),* Volume 9, Issue 1, 2012, (Accepted).

22. S. Karmakar, D. Mukhopadhayay, D. R. Choudhury, "CAvium - Strengthening Trivium using Cellular Automata", *Journal of Cellular Automata*, 2012, (Accepted).

23. S. Kundu, S. Chattopadhyay, I. Sengupta and R. Kapur, "A Diagnosability Metric for Test Set Selection targeting better Fault Detection", In *25th IEEE International Conference on VLSI Design*, 2012.

24. S. Mitra, A. Banerjee and P. Dasgupta, "Formal Methods for Ranking Counterexamples Through Assumption Mining", In *Design Automation and Test in Europe (DATE)*, 2012, (Accepted).

25. S. Mitra, P. Ghosh and P. Dasgupta, "Verification by parts: Reusing Component Invariant Checking Results", *IET Computers & Digital Techniques Journal*, Volume 6, Issue 1, pp. 19-32, January 2012.

26. S. Saha, N. Ganguly and A. Mukherjee, "Information Dissemination Dynamics in Delay Tolerant Network: A Bipartite Network Approach", In *Third International Workshop on Mobile Opportunistic Networks ACM MobiOpp*, Zurich, Switzerland, pp. 15-16, March 2012.

27. S. Saha, N. Ganguly and A. Mukherjee, "Understanding Information Dissemination Dynamics in Delay Tolerant Networks using Theory of Bipartite Networks", In *the Fourth International Conference on Communication Systems and Networks (COMSNETS)*, *PhD Forum*, Bangalore, India, pp. 3-7, January 2012.

2011

1. A. Ain, D. Pal, P. Dasgupta, S. Mukhopadhyay, R. Mukhopadhyay and J. Gough, Chassis: A Platform for Verifying PMU Integration Using Auto-generated Behavioral Models, In ACM Transactions on Design Automation of Electronic Systems (TODAES), Article-33, Volume-16, Issue-3, June 2011.

2. A. Komuravelli, S. Mitra, A. Banerjee and P. Dasgupta, "Backward Reasoning with Formal Properties: A methodology for bug isolation on simulation traces", In *Asian Test Symposium (ATS)*, New Delhi, pp. 238 -243, November 2011.

3. A. Srivastava, B. Mitra, F. Peruani, and N. Ganguly, "Attacks on correlated peer-to-peer networks: An analytical study," In *Computer Communications Workshops, IEEE INFOCOM Workshop*, Shanghai, China, 2011.

4. B. Das, S. Mandal, and P. Mitra, "Bengali speech corpus for continuous automatic speech recognition system," In *International Conference on Speech Database and Assessments (Oriental COCOSDA)*, pp. 51-55, October 2011.

5. C. Karfa, K. Banerjee, D. Sarkar, and C. Mandal, "Equivalence Checking of Array-Intensive Programs", In *IEEE Computer Society Annual Symposium on VLSI (ISVLSI)*, pp. 156-161, 2011.

6. C. Karfa, D. Sarkar, C Mandal, "Verification of Register Transfer Level Low Power Transformations", in *IEEE Computer Society Annual Symposium on VLSI (ISVLSI)*, pp. 313-314, July 4-6, Chennai, India, 2011.

7. C. Rebeiro and D. Mukhopadhyay, "Cryptanalysis of CLEFIA using Differential Methods with Cache Trace Patterns", In *the Proceedings of the Topics in Cryptology - CT-RSA 2011* - The Cryptographers' Track at the RSA Conference 2011, Lecture Notes in Computer Science, San Francisco, CA, USA, pp. 89-103, 2011.

8. C. Rebeiro, R. Poddar, A. Datta, and D. Mukhopadhyay, "An Enhanced Differential Cache Attack on CLEFIA for Large Cache Lines", In *the Proceedings of the 12th International Conference on Cryptology in India*, INDOCRYPT, Lecture Notes in Computer Science, Chennai, India, pp. 58-75, 2011.

9. C. Rebeiro, S. S. Roy, D. S. Reddy, and D. Mukhopadhyay, "Revisiting the Itoh-Tsujii Inversion Algorithm for FPGA Platforms", *IEEE Transactions on Very Large Scale Integration Systems*, Volume 19, Issue 8, pp 1508-1512, August 2011.

10. D. P. Dogra, A. K. Majumdar, S. Sural, J. Mukhopadhyay, S. Mukherjee, A. Singh. "Automatic Adductors Angle Measurement for Neurological Assessment of Post-

neonatal Infants during Follow Up”. In Proceedings of the 4th *International Conference on Pattern Recognition and Machine Intelligence (PReMI),* Moscow, vol. 6744, pp. 160-166, 2011.

11. 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 (IJEHMC),* IGI Global Publication, vol. 2, no. 2, pp. 1-13, April 2011.

12. M. S. Desarkar, “*Multi-parameter Auctions: Truthfulness Conditions and Applications to Page Allocation in Distributed Shared Memory Multiprocessors*”, LAMBERT Academic Publishers, June 2011, (ISBN: 978-3844396089).

13. M. S. Desarkar, R. Joshi, S. Sarkar, “Displacement Based Unsupervised Metric for Evaluating Rank Aggregation”, In *4th International Conference on Pattern Recognition and Machine Intelligence (PReMI),* Moscow, 2011.

14. M. Shirole, A. Suthar, and R. Kumar, “Generation of Improved Test Cases from UML State Diagram using Genetic Algorithm", In *Proceedings of the 4th India Software Engineering Conference, ISEC '11,* pp. 125-134, 2011.

15. M. Tunstall, D. Mukhopadhyay,S. S. Ali, “Differential Fault Analysis of the Advanced Encryption Standard Using a Single Fault”, In *Workshop in Information Security Theory and Practice* (*WISTP 2011),* pp. 224-233, 2011.

16. N. Mishra, R. Saha Roy, N. Ganguly, S. Laxman and M. Choudhury, “Unsupervised Query Segmentation Using only Query Logs”, in Posters of the 20th International World Wide Web Conference 2011 (WWW 2011), Hyderabad, India, pp: 91 – 92 , March 2011 (companion)

17. P. Ghosh, A. Hazra, N. Bhilegaonkar, P. Dasgupta, C. Mandal and K. Paul; “POWER-SIM : An SOC Simulator for Estimating Power Profiles of Mobile Workloads”, In *the International Symposium on Electronic System Design (ISED),* pp. 273-278, December 2011.

18. R. Poddar, A. Datta, and C. Rebeiro, "A Cache Trace Attack on CAMELLIA", In *the Proceedings of First International Conference on Security Aspects in Information Technology, InfoSecHiComNet 2011*, Lecture Notes in Computer Science, Haldia, India, pp. 141-156, 7011, 2011.

19. R. Saha Roy, N. Ganguly, M. Choudhury and N. K. Singh, “Complex Network Analysis Reveals Kernel-Periphery Structure in Web Search Queries”, In *Proceedings of the 2nd International Association for Computing Machinery Special Interest Group on*

Information Retrieval (ACM SIGIR) Workshop on *Query Representation and Understanding 2011 (QRU 2011)*, Beijing, China, pp. 5-8, 28 July 2011.

20. S. Bag and G. Harit, "A novel topographic feature extraction method for Indian character images", In *International Conference on Computer Science and Information Technology (CCSIT)*, Bangalore, India, 2011.

21. S. Bag and G. Harit, "An improved contour-based thinning method for character images", *Pattern Recognition Letters*, Volume 32, Issue 14, 2011.

22. S. Bag and G. Harit, "Skeletonizing character images using a modified medial axis-based strategy", *International Journal of Pattern Recognition and Artificial Intelligence* (IJPRAI), Volume 25, Issue 7, 2011.

23. S. Bag and G. Harit, "Topographic feature extraction for Bengali and Hindi character images", *Signal and Image Processing: An International Journal* (SIPIJ), Volume 2, Issue 2, 2011.

24. S. Bag, G. Harit, and P. Bhowmick, "Topological features for recognizing printed and handwritten Bangla characters", In *Joint Workshop on Multilingual OCR and Analytics for Noisy Unstructured Text Data (J-MOCR-AND) at International Conference on Document Analysis and Recognition (ICDAR)*, Beijing, China, 2011.

25. S. Bag, P. Bhowmick, and G. Harit, "Recognition of Bengali handwritten characters using skeletal convexity and dynamic programming", In *International Conference on Emerging Applications of Information Technology (EAIT),* Kolkata, India, 2011.

26. S. Bag, P. Bhowmick, G. Harit, and A. Biswas, "Character segmentation of handwritten Bengali text by vertex characterization of isothetic covers", In *National Conference on Computer Vision, Pattern Recognition, Image Processing and Graphics (NCVPRIPG)*, Hubli, India, 2011.

27. S. Bag, P. Bhowmick, P. Behera, and G. Harit, "Robust binarization of degraded documents using adaptive-cum-interpolative thresholding in a multi-scale framework", In *International Conference on Image Information Processing (ICIIP),* Shimla, India, 2011.

28. S. Bhattacharyya, A. Roy, D. P. Dogra, A. Biswas, J. Mukhopadhyay, A. K. Majumdar, B. Majumdar, S. Mukherjee, and A. Singh, "Summarization of Neonatal Video EEG for Seizure and Artifact Detection", In *National Conference on Computer Vision, Pattern Recognition, Image Processing and Graphics* (*NCVPRIPG)*, pp. 134-137, 2011.

29. S. Bhattacharyya, J. Mukhopadhyay, A. K. Majumdar, B. Majumdar, A. K. Singh and C. Saha, "Automated Burst Detection in Neonatal EEG". In *International Conference on Bio-inspired System and Signal Processing (BIOSIG)*, Rome, Italy, pp. 15-21, 2011.

30. S. Bhattacharyya, S. Ghoshal, A. Biswas, J. Mukhopadhyay, A. K. Majumdar, B. Majumdar, S. Mukherjee and A. K. Singh, "Automatic Sleep Spindle Detection in Raw EEG Signal of Newborn Babies", In *International Conference on Electronics Computer Technology (ICECT),* Kanyakumari, India, pp. 73-77, 2011.

31. S. Dey, D. Sarkar, A. Basu, "A Kleene Algebra of TSM actors", In IEEE Embedded Systems Letters, Volume 3, Issue 1, pp. 28-31, March 2011.

32. S. Dey, P. Rokkam, A. Basu, "Modeling and Analysis of Embedded Multimedia Applications using Colored PetriNets", In International Journal of Modeling, Simulation, and Scientific Computing (World Scientific), Vol. 2, Issue: 2, pp. 169-193, 2011.

33. S. G. Vadlamudi, P. P. Chakrabarti, Dipankar Das, and Purnendu Sinha, "A Framework for Early Stage Quality-Fault Tolerance Analysis of Embedded Control Systems", In *IEEE/IFIP 41st International Conference on Dependable Systems & Networks (DSN),* Hong Kong, China, pp.315-322, 27-30 June 2011.

34. S. G. Vadlamudi, S. Aine, and P. P. Chakrabarti, "MAWA*—A Memory-Bounded Anytime Heuristic-Search Algorithm", *IEEE Transactions on Systems, Man, and Cybernetics*, Part B: Cybernetics, Volume 41, Issue 3, pp.725-735, June 2011.

35. S. Ghosh and A. Das, "An improvement of linearization-based algebraic attacks", In *International Conference on Security Aspects in Information Technology, High-Performance Computing and Networking*, Haldia, 2011.

36. S. Ghosh, A. Srivastava and N. Ganguly, "Assessing the Effects of a Soft Cut-off in the Twitter Social Network", In *IFIP Networking Conference 2011*, Valencia, Spain, May 2011.

37. 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.

38. 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.

39. S. Karmakar and D. R. Chowdhury, "NOCAS: A Cellular Automata Based Strem Cipher", In *Automata*, 2011.

40. S. Karmakar, D. R. Chowdhury, "Fault Analysis of Grain-128 by Targeting NFSR", In *Annual International Conference on the Theory and Applications of Cryptology (AfricaCrypt),* July, 2011.

41. S. Kundu, S. Chattopadhyay, I. Sengupta and R. Kapur, “Multiple fault diagnosis based on multiple fault simulation using Particle Swarm Optimization”, In *24th IEEE International Conference on VLSI Design*, 2011.

42. S. Mandal, B. Das, P. Mitra, A. Basu, “An Optimum Text Selection Method to Develop Bengali Speech Corpus for Bengali Phone Recognition System”, In *International Conference on Asian Language Processing (IALP),* Penang, Malaysia, 2011.

43. S. Mandal, B. Das, P. Mitra, and A. Basu, “A Robust Bengali Continuous Speech Recognition System Using Triphone clustering and Trigram Language Model”, In *4th International Conference on Contemporary Computing,* Noida, 2011.

44. S. Mohanta, G. Vinod, and R. Mall, “A Technique for Early Prediction of Software Reliability Based on Design Metrics”. *International Journal of System Assurance Engineering and Management*, 2011, (Accepted).

45. S. Mukherjee, P. Dasgupta, S. Mukhopadhyay, “Auxiliary Specifications for Context-Sensitive Monitoring of AMS Assertions”. *IEEE Transactions on CAD of Integrated Circuits and Systems* vol. 30, no. 10, pp. 1446-1457, 2011.

46. S. Mukherjee, P. Dasgupta. “Auxiliary State Machines and Auxiliary Functions: Constructs for Extending AMS Assertions”. VLSI Design, pp. 52-57, 2011.

47. S. Pratihar and P. Bhowmick, “On applying the Farey sequence for shape representation in Z 2 “, In Book Chapter, Speech, Image and Language Processing for Human Computer Interaction- Multi-modal Advancements, IGI Global, 2011, (Accepted).

48. S. Pratihar and P. Bhowmick, “Skew correction of engineering drawings by digital-geometric analysis of Farey ranks”, In *International Conference on Image Information Processing (ICIIP),* Shimla, India. IEEE Press, 2011.

49. S. Pyne, K. Ray, A. Pal, “Realization of Power Aware Software Prefetching as a Multi-Objective Optimization Problem”, In *Second IEEE International Conference on Computer and Communication Technology (ICCCT-2011),* Allahabad, India, pp. 253 - 259 , 15-17 September 2011.

50. S. Roy, D. P. Dogra, S. Bhattacharyya, B. Saha, A. Biswas, A. K. Majumdar, J. Mukhopadhyay, B. Majumdar, A. Singh, A. Paria, S. Mukherjee. "A Web Enabled Health Information System with an Application to Neonatal Patient Care Services". In *IEEE 2011 International Workshop on Web Services in Healthcare and Application (WSHA 2011),* 2011.

51. S. S. Ali, D. Mukhopadhyay, “A Differential Fault Analysis on AES Key Schedule Using Single Fault”, In *the 8th Workshop on Fault Diagnosis and Tolerance in Cryptography (FDTC) 2011*, pp. 35-42, 2011.

52. S. S. Ali, D. Mukhopadhyay, “An Improved Differential Fault Analysis on AES-256”, In *the 4th International Conference on Cryptology and Information Security AFRICACRYPT 2011*, pp. 332-347, 2011.

53. S. S. Ali, D. Mukhopadhyay, “Differential Fault Analysis of AES-128 Key Schedule Using a Single Multi-byte Fault”, In *Smart Card Research and Advanced Applications (CARDIS),* pp. 50-64, 2011.

54. S. S. Ali, R. S. Chakraborty, D. Mukhopadhyay, S. Bhunia, “Multi-level attacks: An emerging security concern for cryptographic hardware”, *Design, Automation & Test in Europe (DATE)*, pp: 1176-1179, 2011.

55. S. S. Roy, C. Rebeiro, and D. Mukhopadhyay, "Accelerating Itoh-Tsujii Multiplicative Inversion Algorithm for FPGAs", In *the Proceedings of the 21st edition of the great lakes symposium on Great lakes symposium on VLSI (GLSVLSI)*, Lausanne, Switzerland, ACM, pp. 67-72, 2011.

56. S. S. Roy, C. Rebeiro, and D. Mukhopadhyay, "Generalized High Speed Itoh-Tsujii Multiplicative Inversion Architecture for FPGAs”, *VLSI Journal*, INTEGRATION, Elsevier, November 2011, (Accepted).

57. S. S. Roy, C. Rebeiro, and D. Mukhopadhyay, "Theoretical Modeling of the Itoh-Tsujii Inversion Algorithm for Enhanced Performance on k-LUT based FPGAs", In *the Proceedings of the Design, Automation and Test in Europe Conference and Exhibition (DATE 2011)*, IEEE Computer Society, 2011.

58. S. Saha and R. Kumar, “Bounded-diameter MST instances with hybridization of multi-objective EA”, *An International Journal Computer Applications* Volume 18, Issue 4, PP. 17-25, 2011.

59. S. Saha, G. Baboo, and R. Kumar, “An Efficient EA with multipoint guided crossover for bi-objective graph coloring problem”, In *Fourth International Conference of Contemporary Computing (IC3)*, pp. 135-145, Springer, 2011.

60. S. Saha, R. Kumar, “Improvement of Bounded-diameter MST instances with hybridization of multi-objective EA”, In *the International Conference on Communication, Computing & Security* (*ICCCS) 2011*, ACM, 2011.

CSE

Research Scholars' Day

March 17, 2012

Indian Institute of Technology Kharagpur

भारतीय प्रौद्योगिकी संस्थान खड़गपुर