Research I Foundation

Annual Report 2008-09

Table of Contents:

Research Activities of Faculty Members	
Dr. Anil Seth	3
Prof. R. K. Ghosh	4
Prof. Shashank K Mehta	5
Dr. Piyush P Kurur	6
Prof. TV Prabhakar	7
Prof. Somenath Biswas	9
Dr. Surender Baswana	10
Prof. Sumit Ganguly	11
Prof. Amitabha Mukherjee	13
Research Activities of PhD Students	15
Purushottam Kar	15
G.S. Badrinath	16
Ajitha Shenoy K B	17
Surya Prakash	18
Seetha Ramaiah	19
Umarani Jayaraman	20
Vibha Patel	21
Pawan Kumar Aurora	22
Apurba Sarkar	23
Sagarmoy Dutta	24
Deepanjan Kesh	25
Satyam Sharma	26
Chandan Saha	27
Kiran Kumar	28
Amrita Pathak	29

Research Activities of Faculty Members

Dr. Anil Seth Research Activities



Temporal Logics for Interacting Process Classes

During the above period IIT Kanpur author of the project worked on multistack pushdown systems. A multi-stack pushdown system (mpds) has a finite set of control states and a fixed number of stacks. The transition function of a mpds takes as input its control state and the topmost symbols of each stack and may (nondeterministically) do a push or a pop operation on any stack along with a possible change in control state of mpds. Multi-stack pushdown systems can be used to model a class of programs with recursion and threads. Each thread has its own stack for its procedures calls and communication among threads is through the common finite states of mpds. Bounded phase multi-stack pushdown automata have been studied recently.

In our work we show that parity games over bounded phase multi-stack pushdown systems are effectively solvable and winning strategy in these games can be effectively synthesized. We show some applications of our result, including a new proof of a known result that emptiness problem for bounded phase multi-stack automata is decidable. It may also be noted that showing parity games effectively solvable implies decidability of many temporal logics on the configuration graphs of mpds. We then continued this work for higher order multi-stack pushdown system (hmpds). Higher order pushdown systems (hpds) are a generalization of pushdown systems (pds) in that hpds can have nested stacks, such as stack of stacks.

Order of a hpds depends on the depth of nested stacks allowed by it. Higher order push and pop operations are provided to push a copy of the topmost stack of any order and to pop it. We define higher order multi-stack pushdown systems and show that parity games over bounded phase higher order multi-stack pushdown systems are effectively solvable and winning strategy in these games can be effectively synthesized. This work is non-trivial extension of the order-1 case. It also implies decidability of emptiness problem for bounded phase higher 1 order multi-stack automata and decidability of temporal logics on configuration graphs of hmpds in the same way as for mpds case.

Publications

Following publications resulted from this work.

• Anil Seth.: Games on Multi-Stack Pushdown Systems. In proc. Logical Foundations of Computer Science, 2009. LNCS, vol. 5407, Sergei N. Artëmov, Anil Nerode (eds.), pp. 395-408. Springer, Heidelberg 2009.

• Anil Seth.: Games on Higher Order Multi-Stack Pushdown Systems. To appear in proc. Reachability Problems'09 to be held at Ecole Polytechnique, France, during September 23rd-25th, 2009.

Prof. R.K. Ghosh Research Activities



In the last year's progress report we reported about an ongoing work on energy aware routing protocols for wireless sensor networks (WSN). The student who worked this problem has since completed her thesis. As reported earlier, the work is centered on asymmetric links between sensor nodes. Using a combination of simulation over synthetic WSN deployment space, our routing protocol is able to achieve a much better performance than existing symmetric protocols even without clustering. With clustering our protocol showed an improvement of about 400% over LEACH. We also have formalized a mathematical model for energy aware routing and reworked on the protocols according to this model. However, simulations of modified protocols are still pending. We hope to catch up more on this work in the coming semester.

According to last year's plan, I spent a semester during sept-dec, 2008 at University of Texas at Arlington as a visiting faculty. During this time we worked on sensor localization. This work has now been accepted in the special issue on Networked Sensing, Decision Making and Control of the Journal of Control Theory and Applications (scheduled for August 2009). As follow up of this work we received a proposal from European Aeronautics Defence and Space Company (EADS) for a project on communication localization and navigation using ad hoc position sensing techniques. The agreements concerning this project have now been concluded.

The other thread of investigation carried out during this period is related to the problem of tracking intruders while gathering of sensory data from a non convex area using mobile sensors. This work was accepted in International Conference on Distributed Computing and Network 2009 (ICDN 2009) and appeared in proceedings (LNCS vol 5408).

We have started another work on the idea of using mobile agent technology for dissemination of data in a WSN. An agent software is currently being built by some students.

Prof. Shashank K Mehta Research Activities



During the last year we have worked on the problem of representation of cyclotomic fields. We have shown that "any subfield of any cyclotomic field can be represented by a subalgebra of circulant matrices modulo a suitable matrix. Conversely, the polynomial algebra of any circulant matrix A under modulo q(A) is isomorphic to some subfield of a cyclotomic field, where q(x) is any irreducible factor of the minimal polynomial of A."

Dr. Piyush P. Kurur Research Activities



The complexity of Matrix Multiplication

Group representation on graphs and graph isomorphism

Often group theoretic questions are better answered by looking at representations of groups. A drepresentation (linear representation) of a group G is a homomorphism from the group to the group of invertible d x d matrices. In general, by a representation of G on an object X, we mean a homomorphism from G to the automorphism group of X. Our interest is the following group representability problem: Given a group G and a graph X, decide whether G has a nontrivial representation on X. The main interest was to study it in the context of Graph isomorphism problem.

We have the following preliminary results: We show, for example, that the graph isomorphism problem reduces to representability of abelian groups. In the other direction we show that even for solvable groups the representability on graphs is decidable using a graph isomorphism oracle. The reductions hold true even when the groups are presented as permutation groups. Under these circumstances one might be tempted to conjecture that the problem is equivalent to Graph Isomorphism. However we conjecture that this might not be the case. The non-solvable version of this problem seems to be harder than graph isomorphism. For example, we were able to show that representability of groups on trees, a class of graphs for which isomorphism is decidable in polynomial time, is as hard as checking whether, given an integer n and a group G, the symmetric group S_n has a nontrivial subgroup homomorphic to G, a problem for which no polynomial time algorithm is known.

It would be interesting to pinpoint the complexity of this problem as it seems to be connected to graph isomorphism but is not quite the same. The above result is reported in [1]

References

[1] Sagarmoy Dutt and Piyush P Kurur. Representing groups on graphs. In Proceedings of the 34th International Symposium on Mathematical Founda- tions of Computer Science (MFCS) 2009, To appear.

Prof. T.V. Prabhakar Research Activities



This year we continued the work on trying to get architecture design into a formal framework. We experimented with using Game theory as a tool to design an architecture. Some initial results were published as a poster in ISEC-2009(some slide from this are enclosed as a separate file).

We had earlier worked on an architecture documentation tool called archaware -- Architecture Aware. It is a lightweight architecture documentation tool which is aware of architecture documentation knowledge. This year we enhanced the tool with a decision view(a view that allows design decisions to be documented) and also completely rebuilt the tool from scratch increasing its usability.

A new PhD student, Amrita Tripathi, joined our team to investigate the effect of semantic technologies on architecture. That is, if we have a semantic component like an ontology in the solution stack, what is its effect on quality attributes? One could see that modifiability will increase and perhaps performance will come down. Are there any design patterns that help capture these relationships?

Our collaboration with Vrije University has become strong. We managed to get sponsorship from European Union for participating in an exchange program. Two of our PhD students Kiran Kumar and Amrita Tripathi are currently in Amsterdam interacting with the architecture group there – Vrije University has the largest architecture group in Europe and is very active. Because of this EU funding, which came about from the interactions that were possible from Research-I funding, we did not use the monies given by Research-I.

We also designed and developed a major application <u>www.agropedia.net</u>. This is an attempt to inject semantic and social networking technologies into Indian agriculture. It has been very well received.

Quality Attribute Game: A Game Theory based Technique for Software Architecture Design.

Kiran Kumar (vkirankr@iitk.ac.in), Prof. Prabhakar T.V. (tvp@iitk.ac.in) Indian Institute of Technology, Kanpur



Prof. Somenath Biswas Research Activities



Structural Characterization of Good Instances for Randomized Search Strategies The project is a collaborative effort between the group headed by Benjamin Doerr at Max Planck Institute at Saarbrucken and S Biswas (PI), S Baswana and PP Kurur of IITK. The work started with the visit of Benjamin Doerr, Tobias Friedrich and Frank Neumann in May (Doerr and Neumann for two weeks and Friedrich for about a month). We worked on the analysis of an evolutionary randomized search strategy for the single source shortest path problem. The work continued through electronic contact and resulted in settling an open issue, viz., can a single-objective fitness function be as efficient as the known multi-objective fitness one? We settled the question in the affirmative. [BBDFKN09]

During February-March of 2009, Daniel Johannsen, a senior PhD student of Doerr at MPI, spent about five weeks at IITK. Doerr also made a brief visit to IITK during this period. The issue we are looking at is the exact capability of the technique of potential function in drift analysis for the One Max problem. PP Kurur is spending about three weeks at MPI Saarbrucken in July 2009 to further continue the work.

We have also involved students of our Department in certain aspects of the questions of interest of the project. One MTech thesis completed in 2008 concerned necessary and sufficient conditions on input instances for the Metropolis algorithm to work efficiently. [D08]. Through the work of a UG student done in his CS 397 (independent studies course), and then continued in the summer of 2009, we have settled the question of the effect of the gap between the largest and the second largest stationary distribution probabilities, on the second largest eigen value of the Metropolis algorithm. (The question has relevance also to the protein folding problem.) The work will be reported soon. Two MTech students and one PhD student of IITK are also now working on certain aspects of the topic of the project. References

[BBDFKN09] Surender Baswana, Somenath Biswas, Benjamin Doerr, Tobias Friedrich, Piyush P. Kurur and Frank Neumann; Computing Single Source Shortest Paths using Single-Objective Fitness Functions. In Proc. 2009 ACM Symposium on Foundations of Genetic Algorithms, Orlando.

[D08] Anjana Das; Analysis of Metropolis Algorithm, MTech thesis, 2008. (Available through the Electronic theses and disertations repository of IITK of the Kelkar Library of IITK).

Dr. Surender Baswana Research Activities



Efficient algorithms which solve problems approximately and/or in dynamic scenario : There are a number of applications which require efficient solution of a graph problem in a dynamic environment. Most of these applications are from the area of communication networks, compilers, databases. The goal of a dynamic graph algorithm is to update the solution efficiently after the dynamic changes, rather than having to re-compute it from scratch each time. We worked on the problem of dynamic maintenance of graph spanners. A spanner is a subgraph which is sparse and yet preserves approximate distance between each pair of vertices. In a joint work with Soumojit Sarkar we designed two randomized algorithms for maintaining a sparse spanner of an undirected unweighted graph under insertion and deletion of edges [1]. Our algorithms significantly improve the existing fully dynamic algorithms for graph spanners in centralized as well as distributed environment. The expected size of the spanner maintained at each stage by our algorithms nearly matches the worst case optimal size of a spanner, and the expected time required to process an update (insertion/deletion of an edge) is close to optimal. We also worked on the problem of computing a data structure of small size which can efficiently report approximate distance between any two vertices. This work [2] was carried out jointly with Akshay Gaur, Jayant Upadhyay, and Prof. Sandeep Sen at IIT Delhi. We could design a data structure which occupies sub-quadratic space, takes sub-quadratic time to be processed, and yet, it can answer any approximate distance query in constant time.

Recently we also worked on the problem of designing data structures for shortest paths in the presence of single vertex failure. The problem is relevant for those dynamic networks where vertex failures occur, but quite rarely. Moreover, as new vertices become dead, the old dead vertices may become alive as a result of simultaneous repair going on in the network. As a result there will be only a small number of dead vertices at any given time. The objective is to have a compact data structure using which one can answer answer shortest path query between any two vertices avoiding all the dead vertices in the network. Though the final goal would be to understand the complexity of the problem of shortest path under a given (small) number of vertex failures, handling a single vertex failure appears to be the first natural problem to be analyzed and solved thoroughly on the way to the final goal. In this joint work with Neelesh Khanna, we have designed data structure which can efficiently report approximate shortest paths between any two vertices when any other vertex fails [3]. As an important result, we show that approximate distance oracle of Thorup and Zwick [JACM,52(1):1-24 (2005)] can be made robust to handle any single vertex failure, and that too without any significant increase in their size.

References :[1] Surender Baswana and Soumojit Sarkar; Fully dynamic algorithm for graph spanners with poly-logarithmic update time. In Proc. ACM-SIAM Symposium on Discrete Algorithms (SODA) 2008, San Francisco, pp 1125-1134.

[2] Surender Baswana, Akshay Gaur, Sandep Sen, and Jayant Upadhyay; Distance Oracles for Unweighted Graphs: Breaking the Quadratic Barrier with Constant Additive Error. In Proc. International Colloquim on Automata, Languages, and Programming (ICALP) (1) 2008, Iceland, pp 609-621. [3] Surender Baswana and Neelesh Khanna; Approximate Shortest Path Oracle under vertex Failure. (submitted).

Prof. Sumit Ganguly Research Activities



The data streaming model presents a computational model for a number of applications where data arrives rapidly and continuously and has to be processed in an online fashion. The rate and volume of data arrival makes it impossible to store the input in its entirety in the main memory, or to maintain index structures in secondary storage. Further, many applications of streaming data do not require exact answers to detail queries, rather, it often suffices to obtain a small amount of useful information or aggregate statistics of the stream. A data stream S is viewed as a sequence of updates the form (index, i, v), where, $i \in [n] = \{1, \ldots, n\}$ and v depicts the change in the frequency of i. The frequency of an item i in the stream S is defined as $f_i(S) = \sum_{(index,i,v)} p_{appears in S}^{v}$.

We develop a theory of stream automaton for characterizing deterministic stream computations over integer frequency vectors. We use it to study the problem of deterministically estimating frequencies (i.e., return f_i such that $|f_i - f_i| \le \varepsilon f_i$). This is used to present stronger lower bounds ($\Omega \in \varepsilon^{-2}$) as compared to the known lower bound $\Omega \in \varepsilon^{-1}$) for the problem of deterministically estimating frequencies to within an accuracy of ε [2].

We consider the issue of flexible massive parallelization of frequency dependent tasks, inspired by the popular MAP-REDUCE paradigm. Using the theory of stream automaton, we show that if there is a parallel computation tree with vectors at the leaf and the tree computes a total function of the sum of these vectors, then, every computation tree can be programmed to do the same [3].

This shows that the parallel computation of total function of distributed sum of vectors is fully flexible. We present a novel algorithm for finding frequent items (i.e., return a set of items i such that $f_i \ge \in L_p$ and do not return any item i such that $f_i < (\notin -\phi) L_p$, where, $L_p = (\sum_{i=1}^{n} |f_i|^p)^{v_p}$ and $p = \notin \{1, 2\}$). We consider update data streams, that is, streams that allow arbitrary insertion

and $p \in \{1, 2\}$). We consider update data streams, that is, streams that allow arbitrary insertion and deletion of items. Our algorithm [5] improves upon the space requirement of $O(\phi^{-2})$ of the only existing algorithm of Cormode and Muthukrishnan to $O(\phi^{-1})$ for p = 2 and is experimentally shown to have better precision and recall.

We demonstrate a novel and simple paradigm for designing algorithms for data stream processing using expander graphs [1]. We pose the problem of k-l separator over a data stream that decides whether the number of items i with $f_i \neq 0$ is at most k or is at least l. The k-2k-separator problem for streams with non-negative integer frequency vector is solved using an application of expander graphs. This is then used to present the first space-optimal solution for the k-sparsity problem (i.e., a k-k + 1 separator) for non-negative integer frequency vectors. We consider the problem of hybrid frequency moments over two-dimensional streams. In a two dimensional stream, each stream record is of the form (i, j, v) signifying that $A_{i,j}$ is updated to $A_{i,j} + v$, where, $1 \le i, j \le n$. The p, q hybrid moment is defined as $F_{p,q} = \sum_{j}^{n} =_{1} \sum_{i}^{n} =_{1} (|A_{i,j}|^{p})^{q}$. Estimation of hybrid moments finds applications in network monitoring. We present the first algorithm with space requirement that is poly-logarithmically dependent on the size of the matrix for the range q $\in [0, 1]$ and p $\in [0, 2]$ [4]. We also present an $O(n^{1-1/q})$ space algorithm for q > 1 and p $\in [0, 2]$.

In addition, the author was invited to present a talk and participate in the elite forum Dagsthul Seminar series on Sub-linear Algorithms (Seminar No. 8341) and to present a seminar on stream automaton at the University of Frankfurt, Germany. The papers listed below were presented at the locations listed funded by the Research I foundation. References

[1] S. Ganguly. "Data Stream Algorithms via Expander Graphs". In Proceedings of International Symposium on Algorithms, Automata and Computation (ISAAC), Gold Coast, Australia, pages 52–63, 2008.

[2] S. Ganguly. "Lower bounds for frequency estimation over data streams". In Proceedings of the Computer Science Symposium of Russia (CSR), Springer LNCS 5010, Moscow, Russia, pages 204–215, 2008.

[3] S. Ganguly. "Distributing Frequency-Dependent Data Stream Computations". In Proceedings of Computing: the Australasian Symposium (CATS), Wellington, New Zealand, January 2009.

[4] Sumit Ganguly, Mohit Bansal, and Shruti Dube. "Estimating hybrid frequency moments of data streams". In International Frontiers of Algorithms Workshop, Changsha, Hunan, China, June 2008.

[5] Sumit Ganguly, Abhayendra N. Singh, and Satyam Shankar. "Finding frequent items over general update streams". In International Conference on Scientific and Statistical Database Management (SSDBM), Hong Kong, July 2008.

[6] Sumit Ganguly "Lower bounds for frequency estimation over data streams". Talk presented in Dagsthul Seminar on Sub-linear Algorithms, Dagsthul, Germany, August, 2008.

Prof. Amitabha Mukherjee Research Activities



Developmentally Motivated Models for Learning Ontology from Perception"

The project involves formulating computational models analogous to mental models for perceptual concepts, and how these impact problem-solving tasks such as design. These involve the spatio-temporal signature of actions (called "image-schema" in cognitive psychology). Tightly coupled with perception is the motor actions of the agents (particularly in situations where the camera is on moving platforms.

In 2008-09, new work was done on linking perception to motor activity, while unsupervised ontologies were extended to actions, and implicit models in design were related to manifold learning in high-dimensional spaces. These three areas are detailed below.

Also an international conference is proposed for late 2010/early 2011 on "Cognitive Issues in Language and Perception". Speaker identification is in progress.

A. Linking perception with motor tasks: How a moving camera may map the environment and calibrate its own position based on multiple observations over time.

In human infant cognition, perception is combined with motor activity. In the early phase of this work (reported in 2007-2008), perception was investigated, particularly in identifying the spatio-temporal signature of activities. In the past year, we have begun to explore the combination of motor with perceptual activity by considering mobile robots that obtain images as they move and are able to calibrate themselves merely by moving through certain spaces.

B. Spatial Design: Implicit models of feasible regions may be low-dimensional manifolds in feature space Work on modeling interaction between objects (a key competence in problemsolving tasks such as mechanical design), saw a key step emerge in the past year. As an infant attempts to insert various objects into various cavities, it gradually forms a model of containment. By the age of three months, babies are sensitive to that large objects cannot enter small, open cavities. This involves both perception (one object is occluded) and also motor activity (the object cannot be pushed in). The joint sensorimotor space of containment is high-dimensional and involves many aspects of relative pose between objects. In children as well as adults, this knowledge is often implicit, in that it is not available to consciousness. In recent work, we have been able to show that feasible regions (where a task such as insertion may be successful) are related to lower-dimensional manifolds in this high-dimensional space, which are a possible mapping for these implicit awareness episodes. Over repeated experience, such a model may become the key constituent in a process of symbol discovery. C. Spatial action modeling and learning action ontologies. Earlier, an unsupervised model was trained to learn action schemas from observing 2D and 3D videos. Here these action models (image schemas) were shown to constitute a hierarchy. This involved both non-commutative actions such as {\em chase}. Also, commutative actions like "come-closer" were shown to have non-commutative sub-classes, e.g. or "come-closer-one-object-static".

Personnel involved: Post-phD: Madan Dabeeru Ph.D.: P. Guha M.Tech. S. Daftadar, D. Jogle

SELECTED PUBLICATIONS (2008-2009):

Amitabha Mukerjee, Discovering symbols from interactions easier than explaining interactions via symbols? Workshop on Logic and the Simulation of Reasoning, Pasadena CA Jul 12, 2009

Madan M. Dabeeru, Amitabha Mukerjee, Product Platform selection in Lower-Dimensional Manifold Spaces, DETC conference August 31-Sept 2, 2009, San Diego.

Amitabha Mukerjee and Madan Dabeeru, Symbol emergence in design, Fifth International Workshop on Neural-Symbolic Learning and Reasoning NeSy09, Jul 11, 2009

Amitabha Mukerjee, Madan M. Dabeeru, The birth of symbols in design DETC-DTM Conference, San Diego August 31-Sept 2, 2009.

Akhilesh Kumar Sinha, Prithwijit Guha, Amitabha Mukerjee: Back to the future: Robust foreground extraction with reversed-time background modeling. ICPR, Dec 2008

Prithwijit Guha, Amitabha Mukerjee, and K.S. Venkatesh Occlusion Sequence Mining for Activity Discovery from Surveillance Videos In Pattern Recognition Technologies and Applications: Recent Advances, ed. Brijesh Verma and Michael Blumenstein, Information Science Reference Publishers, Australia, 2009.

G. Satish, Amitabha Mukerjee, Acquiring linguistic argument structure from multimodal input using attentive focus, 7th IEEE International Conference on In Development and Learning, 2008. ICDL 2008. p. 43-48.

Amitabha Mukerjee, Using attentive focus to discover action ontologies from perception Fifth International Workshop on Neural-Symbolic Learning and Reasoning NeSy09, Jul 11, 2009

PhD Students Supported by RIF

Purushottam Kar Research Activities



I joined the CSE/IITK department in July 2008 with a bachelor's degree in Computer Science. Consequently my academic load in the past two semesters has mainly consisted of course-work. However I have got opportunities to look at various research problems by doing projects which formed a part of the requirement of the courses. Of these I give a brief account below:

On Low Distortion Embeddings of Statistical Distance Measures into Low Dimensional Spaces

This was joint work with Manjish Pal done under the guidance of Prof. Arnab Bhattacharyya and was done as a part of the course CS618 - Indexing and Searching Techniques in Databases. The work studied various statistical distance measures such as the Kullback-Leibler divergence, Mahalanobis distance and Bhattacharyya distance from the point of view of embedding high dimensional vectors under these distance measures into low dimensional ones (we also considered embeddings into metric spaces) preserving the inter-vector distances. We provided some positive results in this direction and some negative ones. This work was accepted for publication in the 20th International Conference on Database and Expert Systems Applications (DEXA'09).

Lower Bounds in the Stream Automaton Model :This was joint work with Piyush Srivastava done under the guidance of Prof. Sumit Ganguly and was done as a part of the course CS719 – Introduction to Data Stream Algorithms. The work consisted of proving lower bounds on various data stream problems using the recently proposed Stream Automaton model for data stream computations. We provided lower bound proofs for the k-sparsity testing problem and the problem of deterministic estimation of norms in general streams. Although the results we gave were already known, our proofs which utilized the Streaming Automaton model were simpler than existing proofs and indicated that the model could be extended to give a unified model for proving lower bounds in data streaming problems.

Learning Manifold Classifiers :This was joint work with Manjish Pal done under the guidance of Prof. Harish Karnick and was done as a part of the course CS797 – Special Advanced Topics in Computer Science. In this work we took up the problem of classification in Euclidean spaces and investigated cases where the classification can be done using a smooth manifold. Usually in such cases one applies a non-linear transformation (the Kernel Trick) to transform the problem into one where a hyperplane is able to perform the classification. But the problem of finding such a transformation is not well studied. Our contribution in this work was to develop techniques that allow one to learn similarity measures based on training data that can help the process of identifying the "linearizing" transformation.

Apart from this I have done several other reading projects for courses in Bioinformatics and Automata Theory of which I do not give details here. I was also actively involved with a reading group called SIGTACS (Special Interest Group on Theoretical Aspects of Computer Science) where the group members gave weekly seminars on a plethora of topics. Presently I am trying to gain a deeper understanding of Kernel methods in Machine Learning and aspects of Computational Learning Theory. I am also trying to gain knowledge in the field of Circuit Complexity and am in the process of finding problems to work on in the coming months. I am also preparing for a short informal course on Approximation Algorithms that SIGTACS is planning to offer next semester which some students (including myself) would be "teaching".

G.S. Badrinath Research Activities



The traits like Finger print, Iris, and Face are long been studied and developed to provide a reliable verification system. I have chosen to develop a system based on palmprint. Despite of its features and uniqueness limited work is reported in literature and found the proposed systems has following limitations. The palmprint for enrolling/verifying the user is obtained through constraints (pegs). Hence not it is not suitable for people with medical injuries / physically challenged. Furthermore, the systems

reported are not suitable for real time application, because they could not handle the situations such as, given partial palmprint and asked to identify the user; given image with different size (Resolution) and asked to identify the user. The systems proposed will work for hand images obtained from scanner of definite specifications and definite environment conditions.

I am motivated to develop a palmprint based verification system which performs with high accuracy and incorporating the following:

- 1. Constraint free hand image acquisition for Enrolling and verifying.
- 2. Robust to Occlusion: Able to Identify/verify user with partial palmprint.
- 3. Robust to Scale: Able to Identify/verify user from images with different resolution.
- 4. Robust to Camera used: Able to work with any scanner of any specifications.

The proposed verification system using palmprint acquired hand images from flat bed scanner. The scanned is constraints (Pegs) free. Hence the user with medical injuries and physically challenged can also use the system. A procedure to extract palmprint from the hand images obtained is proposed and found to be robust to rotation and translation of plamprint placement on scanner bed. Global approaches like Zernike moments, and local approaches like Speeded Up Robust Features, and Scale Invariant Feature Transform are used to extract palmprint features. During verification the extracted features of the palmprint are compared using distances like Euclidean and Manhattan distance. Systems using local approaches for palmprint are robust to Scale, Roataion and Occlusion but have low accuracy for large scale database. Systems using global features describing palmprint are found to be performing with high accuracy but not robust to Rotation, Scale or Occlusion.

The proposed system is tested on database collected at Indian Institute of Technology Kanpur which consists of 150 hand images. It is found that the proposed system performs with accuracy 99.98% with False Acceptance Rate less than 0.01%. The system is also tested on large database of 7752 hand images obtained from The Hong Kong Polytechnic University and found to be performing with accuracy more than 99.5%. Hence the system is robust to camera used to obtain the hand images. Proposed systems with various approaches for features are found to outperform the best known system proposed by Zhang et. al. in the literature. The proposed systems take less than 2sec to verify the user. The proposed systems being highly accurate, Speed, Low False acceptance Rate, Robust to Rotation, Robust to Occlusion, and Robust to Scale suggests it application for both High end security and civilian applications.

Ajitha Shenoy K B Research Activities



I completed my course work with 9 CPI. I am working under Dr. Somenath Biswas. We are working in LLL algorithm. Our aim is to either reduce the time complexity of this algorithm or to get new algorithm for finding shortest vector in the lattice. We have proved that finding a vector which is shorter than some specified vector λ is in NP. We found the bound on each variable.

In august I will give my oral comprehensive exam. I am attending summer school in evolutionary algorithm at Finland (from 16 - 22 August). Then I will prepare for SOTA. Our aim is to apply evolutionary algorithm to shortest vector problem and study the results.

Surya Prakash Research Activities



Development of a robust and efficient ear based human recognition system

Ear biometrics has been found to be a good and reliable technique for human recognition in recent years. Since it has been found that it is almost impossible to find two ears with all the parts identical, it has gained much pace. My research work aims to develop a robust and efficient human recognition method using ear information.

Any ear based recognition procedure follows three steps viz. ear localization from input image, feature extraction from the detected ear, and matching using computed features. My current research work deals with the first part of the recognition process, i.e. ear localization, and focuses on the automation of it which is very essential to automate an ear based recognition process. Work related to this research appears in following research papers.

Related Published Papers:

- 1. Surya Prakash, Umarani J and Phalguni Gupta, "Connected Component based Technique for Automatic Ear Detection", 16th IEEE International Conference on Image Processing (ICIP 2009), Cairo, Egypt, Nov 2009.
- 2. Surya Prakash, Umarani J and Phalguni Gupta, "Ear Localization using Hierarchical Clustering", Proc. of SPIE Int'l Defense Security and Sensing conference (Biometric Technology for Human Identification VI), Vol. 7306, Orlando, Florida, April 2009.
- 3. Surya Prakash, Umarani J and Phalguni Gupta, "A Skin-Color and Template Based Technique for Automatic Ear Detection", Proc. of ICAPR 2009, IEEE Computer Society, pp. 213-216, Kolkata, India, Feb 2009.
- Surya Prakash, Umarani J and Phalguni Gupta, "Ear Localization from Side Face Images using Distance Transform and Template Matching", Proc. of IEEE Int'l Workshop on Image Proc. Theory, Tools and Applications, IPTA 2008, pp. 1-8, Sousse, Tunisia, Nov 2008.

Seetha Ramaiah Research Activities



I have joined here during the December '08 semester and I am just done with my course work. I still have to identify the problem to be worked upon. The area of interest is **autopoiesis**. It mainly deals with making the software self-organizing. These days I am reading a few papers in this area. And in the month of September I will probably go to the London School of Economics for research collaboration with Prof Paolo Dini and his research team. I am hoping to know the state of the art in this area through the interactions with those people.

Umarani Jayaraman Research Activities



Efficient search and retrieval techniques in multi-modal biometric database:

Biometric system provides an automated method to verify or to identify an individual based on unique behavioral or physiological characteristics. The task of the authentication module in a biometric system is to recognize a subject either by identification of one person among many, or verification that a person's biometric matches a claimed identity. In case of identification, for given query image it has to compare the whole database to declare its identity.

In case of security checks at airports and border crossing, the biometric database is very large. Performing an exhaustive search in a database involving billion of comparisons will be computationally expensive. If an effective search method is designed for a biometric database, it will reduce the number of comparisons on biometric database. Thus, the computational complexity in searching a given query image in a very large database will be reduced significantly. In other words, the indexing method selects a small subset of images in the database from which the feature matching algorithm determines the correct match.

Related Papers:

- 1. Umarani J, Surya Prakash and Phalguni Gupta, An Efficient Technique for Indexing Multimodal Biometric Databases, International Journal of Biometrics, and March 2009 (accepted).
- Umarani J, Surya Prakash and Phalguni Gupta, "Indexing Multimodal Biometric Databases Using Kd-tree with Feature Level Fusion", Proc. of ICISS 08, LNCS 5352, pp. 221-234, Hyderabad, India, Dec 2008.
- Umarani J, Surya Prakash, Dev Dutt and Phalguni Gupta, "An Indexing Technique for Biometric Database", Proc. of ICWAPR 08, IEEE Computer Society, Vol. 2, pp. 758-763, Hong Kong, Aug 2008.

Vibha Patel Research Activities



I joined the Ph.D programme of the institute in July-2008. Since this was my first year of the programme I took four courses in the first semester and finished my course work. The courses include Parallel Execution of Programs (CS629), Indexing and Searching Techniques in Databases (CS618), Advanced Compiler Optimizations (CS738) and Advanced Topics in Algorithms and Data Structures (CS647).

We did project on "*Comparison of PCA and Subspace LDA for Face Recognition*" as part of Indexing and Searching Techniques in Databases course. We used Principal Component Analysis and Subspace Linear Discriminant Analysis to extract significant features with good discrimination ability. This will reduce the dimensions of the face image used for online comparison, which in turn will improve the response time of the system. We also did a project on "Study of Widening Operator for Program verification" as part of Advanced Compiler Optimizations course. It was a study project where we did a literature survey on static program analysis using Abstract interpretation. It is basically concerned with automatically discovering properties of a program that hold for all possible execution paths of the program.

In second semester I faced my comprehensive viva exam in the month of February. I successfully completed the exam in first attempt. I was given a paper on "Serializability violation detector for shared memory server programs" to be presented during the viva exam. Since then I am working on my thesis work which is on High performance computing using Graphics processing Unit.

I visited nVidia, Pune in month of March, where I had an interaction with Compiler group and Multimedia group. I got a broad idea on various areas of research, through the interaction with both the group. I also went to IIT-Madras in the first week of June to attend a workshop on Optimizing Performance of Parallel Programs on Emerging Multi-Core Processors & GPUs which was jointly conducted by CDAC and IIT-Madras. The topics covered in the workshop includes Performance Enhancement through Software Multi-Threading (Pthreads, OpenMP, Intel TBB; MPI); Multi-Core I/O; Benchmarks; Transactional memory on Multi-Core Processors; Programming and Performance Issues on GPGPUs & GPU Computing.I am preparing for the State of the Art seminar which is like to be held in a month or two. I referred several papers related to my thesis. I am listing few of them below.

- [1] Automatic data movement and computation mapping for multi-level parallel architectures with explicitly managed memories, Proceedings of the 13th ACM SIGPLAN Symposium on Principles and practice of parallel programming, Muthu Manikandan Baskaran et. Al.
- [2] A Practical Algorithm for Exact Array Dependence Analysis, Communications of the ACM, William Pugh
- [3] The Power Test for Data dependence, IEEE transactions on parallel and Distributed systems, Michael Wolfe, Chau-Wen Tseng.
- [4] A New Practical Array Data Dependence Analysis for Parallelizing Compilers, Proceedings of the Innovative Architecture for Future Generation High-Performance Processors and Systems, Masaaki MINEO, Tetsutaro UEHARA, Shoichi SAITO, Yoshitoshi KUNIEDA

Pawan Kumar Aurora Research Activities



As part of the research project for the Data Stream Algorithms course, I designed a randomized algorithm for the maximal matching problem (in the streaming model) that takes on expectaction a constant number of passes and constant amount of space for a bounded degree graph (the degree is assumed to be a constant). It was a simple extension to the FOCS 2008 paper by Nguyen and Onak titled Constant-Time Approximation Algorithms via Local Improvements. I also studied the d-dimensional knapsack problem (in the data streaming model) on which my supervisor Prof. Sumit Ganguly was working. I ran a few experiments on an algorithm designed by Prof. Ganguly for estimating the frequency moment F4 of a data stream.

For my comprehensive exam I read and presented two classic papers on hashing. The first one due to Fredman and others dealt with the problem of static hashing wherein the set to be hashed is known in advance whereas the second one due to Rasmus Pagh allowed insert and delete operations to be intermingled with find queries. Both the hashing techniques achieve worst case constant lookup time.

Since the past few months I have been working on the problem of Local Clustering under the supervision of Prof. Sumit Ganguly. We are trying to improve upon the results obtained by Spielman-Teng (ST) and later by Anderson-Chung- Lang (ACL). ST used the idea of truncated random walk to obtain a local clustering algorithm whereas ACL used the idea of pagerank. Both the methods start with a single vertex having the entire distribution. Prof. Ganguly has shown that the cluster obtained by these methods can be improved by doing a multi-source random walk where the initial distribution is stationary within the starting cluster. He has further shown that if we can obtain a set of vertices of significant overlap with a good cluster, then a multi-source random walk from that set can obtain a cluster of conductance close to that of the good cluster. We are presently working on the problem of finding such a set using the idea of BFS.

In future we might look at the area of network evolution where we attempt to answer questions like how do real graphs evolve over time? It is an interesting area but poses several engineering challenges like collecting and managing real data.

Apurba Sarkar Research Activities



I have joined the department as a QIP PhD student in the even semester(January) 2009 and taken the following courses as my course work.

- 1. CS-651 Concurrent Data-Structure and Algorithm,
- 2. CS-641 Modern Cryptology,
- 3. CS-743 Advanced Graph Algorithm,
- 4. CS-797 Special Topics in Computer Science.

In CS-797 I studied several security services provided by cryptography and also studied public Key Cryptography Standards for cryptographic primitives, encryption schemes and signature with appendix under the supervision of Prof. Rajat Moona.

Currently I am preparing for comprehensive examination. My adviser is Prof. Rajat Moona. My thesis is aimed at developing a security architecture framework for defining the security needs of a computing application, validating the security needs, modeling and verification of the implementations

Sagarmoy Dutta Research Activities



My research area is about quantum computing and complexity theory. In the begining of academic year 2008-2009 I was working under Aswin Nayak in Institute of Quantum Computing, University of Waterloo, Canada as part of Graduate Student Exchange Programme funded by Canadian government. The research topic was to characterise quantum random walk on a non symmetric markov chain which can find a marked state faster than classical random walk, much in the same way that has been done for symmetric chains to obtain a quadratic speedup. Surprisingly we came up with a family of non-symmetric but reversible chains parametrised by a quantity e, on which quantum walk for reversible chains would take arbitraty long time as e approachs zero whereas classical walk will take linear time.

After coming back to IITK campus on September '08 I first started working on a problem related to matrix multiplication under my guide Piyush P Kurur which did not yield any fruitful result. Then we took up the following problem. Given a group in table form and a graph how hard it is to find whether there is any nontrivial homomorphism form the given group to the automorphism group of the graph. We formulate this as representing groups on graphs and show that with respect to polynomial time turing reducibility, both abelian and solvable group representability are all equivalent to graph isomorphism, even when the group is presented as a permutation group via generators. On the other hand, the representability problem for general groups on trees is equivalent to checking, given a group G and n, whether a nontrivial homomorphism from G to the complete symmetric group of n symbols exists. There does not seem to be a polynomial time algorithm for this problem, in spite of the fact that tree isomorphism has polynomial time algorithm. The results are written in a paper named "Representing groups on graphs" which is accepted in 34th International Symposium on Mathematical Foundations of Computer Science for publication.

Currently we are trying to find some complexity bounds for the general group representation problem on graphs. In particular trying to find both classical and quantum interactive proofs. Also we have started working on a problem related to quantum error correction code. We want to explore whether it is possible to find QECC that correct classical error on quantum data more efficiently than classical codes.

Deepanjan Kesh Research Activities



Thesis Supervisor : Prof. Shashank Mehta

Throughout the academic year, I have been studying toric ideals and their applications. Along side, I also have been looking at some problems that I can attempt and hope to solve.

In the simplest of terms, toric ideals are prime, binomial ideals. They are generated by the kernels of certain kind of linear maps. Toric ideals are important because they find widespread applications in problems, such as solving integer programs and finding out primitive partition identities.

Finding out the generators of a toric ideal is a very important problem to tackle and has obvious uses in this area. The best algorithm for computing generators of toric ideals involves computation of n grobner bases in graded reverse lexicographic order, where \$n\$ is the number of variables in the ring containing thetoric ideal. We had hoped to improve on the algorithm by showing that we need to compute only constant number of grobner bases in lexicographic ordering. After spending considerable amount of time on the problem, we could neither find a suitable proof, nor could find a counterexample refuting our hypothesis.

But with the knowledge that the complexity of grobner basis computation is a strong function of the number of variables involved in the ideal, we could come up with an algorithm where, though we are still computing n grobner bases, the bases are being computed incrementally in one, two, ..., n variables.

Satyam Sharma Research Activities



I am a first year Ph.D. student working in the area of security. My other areas of interest are data structures and algorithms, computer architecture, operating systems and computer networks. My thesis advisor is Prof. Rajat Moona.

In the past one year, I have worked on projects dealing with systems security and applied cryptography (such as the design and development of a network accessible encrypting file system) as well as read books and research papers in the area of cryptology (on topics such as formal analysis of authentication protocols and the cryptanalysis of block ciphers and public key cryptosystems).

I am presently reading papers to decide upon an interesting and relevant thesis topic and find suitable open problems. In this regard, I have read material on the foundations of cryptography and protocols for special applications such as electronic cash. Although I have completed the requisite coursework, I plan to attend additional courses in the coming year to build a stronger base in the subjects of probability theory, computational complexity and number theory that are vital to undertake any non-trivial research in my primary area of interest.

Chandan Saha Research Activities



Last year, I have been mostly devoting my research time to the study of a problem known as 'Polynomial Identity Testing' (PIT) in complexity theory. The focus of this study has been to find an efficient deterministic algorithm to solve PIT, which is an open problem for quite some time. This problem is linked with other very fundamental problems in complexity theory and has also been successfully used in designing efficient algorithms for other long-standing problems.

Till date, several approaches to solve the PIT problem have been proposed and they are being deeply studied. But none so far has been totally successful in solving the genearal PIT problem, or even an important special case of the PIT problem known as `PIT for depth-3 circuits'. I, along with my collaborators Ramprasad Saptharishi and Nitin Saxena, have been investigating this later problem last year.

As a result of our study, we could show a 'slightly surprising' equivalence between the two problems, 'PIT for depth-3 circuits' and 'PIT for depth-2 circuits over 2 x 2 matrices'. Such a connection opened up a promising possibility of using structural results from abstract algebra to solve the PIT for depth-3 circuits. Indeed, we could successfully exploit such structural results to solve the PIT problem for depth-2 circuits over a 'large class' of commutative algebras. Further, we have shown that if our approach can be extended to depth-2 circuits over commutative algebras of `higher' dimensions then `PIT for depth-3 circuits' can also be solved efficiently. We have also shown that, `depth-2 circuits over 2 x 2 matrices' are strictly weaker than `depth-3 circuits' as computational models, although (surprisingly) the PIT problems on these models are equivalent. The work is at present submitted to a conference.

Beside this work, I have been writing a survey article, with my supervisor Manindra Agrawal, on 'Techniques Used in Algebraic and Number Theoretic Algorithms'. We expect to submit this article sometime soon.

Kiran Kumar Research Activities



Formalized Mechanisms for Architecture Development Environment:

In this project, we address the automation aspects of Architecture design based on the existing theories for architecture design such as Attribute driven design, Decision view, Application independent knowledge etc.

Problem statement: Identifying some challenging tasks in the architecture development life-cycle for which formalized methodologies are missing, designing formalized methodologies to enable automation of these tasks, and demonstrating well-known processes that use these tasks.

Tasks considered for which formalization solutions are required

- 1. Pattern relationship analysis.
- 2. Architecture-solution alternative Explorer.
- 3. Optimal architecture design decision identification.
- 4. Transforming Decision view to other views.
- 5. Controlling selfishness behavior of stakeholders.
- 6. Architecture risk identification.
- 7. Module design evaluation.
- 8. Architecture refining scenarios elicitation.

We have been working on these issues and have some preliminary results [1], the formalized mechanisms designed for some of the sub-problems are listed below:

Architecture-solution alternative Explorer.	Architecture-primitive-kernel knowledge
	model
Optimal architecture design decision identification.	Quality Attribute Game
Controlling selfishness behavior of stakeholders.	Incentive compatible CBAM
Architecture risk identification.	Side-effect aggregation model.

To improve maturity of our mechanisms, we are now working on integrating our mechanisms with others' tools which achieve enterprise level requirements. For this purpose, we are participating in the GRIFFIN project.

[1]: Kiran Kumar, Prabhakar TV, "Quality attribute game: a game theory based technique for software architecture design", 2nd India Software Engineering Conference, 2009.

Amrita Pathak Research Activities



Software Architecture and Semantic Components: Semantic artifacts like Ontologies are becoming common components in current day systems. Typically these components try to capture business rules and are in the form of SBVR encodings, OWL files, logic predicates etc. A simple example for this is encoding the access privileges as a taxonomy of roles and access rights. Presence of such components poses interesting challenges to the Software Architect. For example, they increase modifiability but perhaps impact performance if invoked at run time.

We would like to explore the relationship between such semantic components and quality attribute response of an architecture design. We would like to explore ideas like tactics and design patterns that occur when a semantic component is in place – like "break semantic dependency". At the end we hope to come up with guidelines on how to incorporate semantic components in an architecture design.

Our problem basically investigates the impact of semantic technologies on software architecture.