

# Poster and Demo Abstracts

## Boston University IAP 2007

### Databases

#### **Indexing Hot Motion Paths with Guarantees (Unpublished Work in Progress)** (Poster)

Michalis Potamias, joint work with Dimitris Sacharidis (NTUA), Kostas Patroumpas (NTUA), Manolis Terrovitis (NTUA), Verena Kantere (NTUA), Kyriakos Mouratidis (SMU) and Timos Sellis (NTUA)

Supervised By: George Kollios

Useful location-based services have become very popular due to advances in location sensing technology. We propose a framework for online discovery and maintenance of important motion paths over a sliding window. The extracted motion paths may be utilized in many ways, such as position prediction and pattern discovery. Our setting is a distributed environment, where each location-aware moving object is capable of reporting its position to a central coordinator. We aim at effective motion-path discovery, while minimizing communication cost. We achieve this by shedding trajectory compression overhead to each object in order to reduce the frequency of location updates. The coordinator maintains an index of the motion-paths by efficiently grouping common paths that objects crossed. Our techniques take into consideration the inherent inaccuracy of the location readings, and provide discrepancy guarantees in the discovered motion paths.

---

#### **Efficient Subsequence Matching Using Embeddings** (Poster)

Panagiotis Papapetrou, Vassilis Athitsos, Michalis Potamias

Supervised By: George Kollios

A method is introduced for efficient subsequence matching, with applications to large time-series and biological sequence datasets. The proposed method can be applied with sequence similarity measures that are based on dynamic programming, such as dynamic time warping (DTW) and the edit distance. Our method constructs an embedding that maps each query sequence into a single vector, and each database sequence into an equally long sequence of vectors. There is a one-to-one correspondence between each such vector and a position in the database sequence. Comparing the embedding of the query with those vectors is used to efficiently identify relatively few areas of interest in the database sequence. Those areas of interest are then fully explored using the exact DP-based subsequence matching algorithm.

---

### Image and Video Computing

#### **Switching Dynamic Global Coordination Model: Manifold Learning for High Dimensional Time Series Data** (Poster)

Rui Li and Taipeng Tian

Supervised By: Stan Sclaroff

This work addresses the problem of constructing a low-dimensional manifold from high-dimensional time series. The main contribution of the proposed approach lies in overcoming the difficulties in modeling the interactions between complex dynamics and the non-linear dimensionality reduction when sculpting the low-dimensional manifold. The proposed approach is able to handle large training datasets, model complex dynamical behavior and produce model parameters for the classification task. In the proposed solution, (1) multiple linear fact analyzers

are used reduce the dimensionality, and (2) multiple linear dynamical models to describe the complex dynamical behavior, of the high-dimensional time series data. The interactions among these linear models during the non-linear manifold learning process are captured by the interactions in a mixed-state dynamic Bayesian network. Variational algorithms are used for the inference and parameter learning of the proposed model. The proposed approach along with competing approaches are tested and compared on the dimensionality reduction and reconstruction of synthetic time series data, dynamic texture synthesis and human motion synthesis, classification and tracking. Quantitative and qualitative evaluation results obtained substantiate the claims of efficient training with large datasets and accurate modeling of complex dynamical behavior (which naturally leads to more accurate classification results).

---

### **Human-Computer Interaction for Universal Access with Multiple Camera Active Appearance Model 3D Face Tracking (Poster)**

John Magee

Supervised By: Margrit Betke

People with severe paralysis typically cannot use traditional human-computer interface devices. Those with limited or no motor control below the neck have been able to use buttons or switches pressed with their heads as a way to control wheelchairs or interact with computers. Computer vision-based interaction systems have more recently enabled other modes of interaction, such as controlling a mouse pointer by tracking a user's head movements. This work builds upon the success of these systems by incorporating new advances in 3D head tracking. Head tracking is typically accomplished with monocular cameras, but the use of multiple cameras may increase the usefulness and accuracy of Human Computer Interaction (HCI) systems for Universal Access (UA). Several methods to analyze and advance Active Appearance Model (AAM) fitting from views of multiple cameras are proposed. The difficulty of working with uncalibrated cameras can be overcome via a variety of methods to estimate the relative orientation between cameras. Existing HCI applications may show improvement with a system that uses multiple cameras, for example, to improve accuracy of feature tracking for a mouse substitute interface, or to deal with a rotated or partially occluded face. A multi-camera mouse-substitution interface will be developed based on this research. In the future, this work will also have applications to face tracking for video surveillance and linguistic analysis of American Sign Language video sequences.

---

### **Parameter Sensitive Detectors (Poster)**

Quan Yuan, Ashwin Thangali and Vitaly Ablavsky

Supervised By: Stan Sclaroff

Object detection can be challenging when the object class exhibits large variations. One commonly-used strategy is to first partition the space of possible object variations and then train separate classifiers for each portion. However, with continuous spaces the partitions tend to be arbitrary since there are no natural boundaries. For example, consider the continuous range of human body poses. In this paper, a new formulation is proposed, where the detectors themselves are associated with continuous parameters, and reside in a continuous function space. There are two advantages of this strategy. First, a partitioning of the parameter space is not needed; the detectors are in a continuous space. Second, the underlying parameters for object variations can be learned from training data in an unsupervised manner. Experiments in profile face detection, hand shape detection, and pedestrian detection demonstrate the advantages of this new approach over past techniques.

---

### **Geometric Coherent Cues for 2D Human Motion Tracking (Poster)**

Tai-Peng Tian, Rui Li

Supervised By: Stan Sclaroff

The goal of this work is to track 2D human body configurations in videos filmed by a moving camera. Without assuming any prior knowledge on the dynamics of the human motion or camera, most robust trackers rely heavily on appearance cues to localize and assemble body parts in each frame. We propose a new set of cues, based on geometry coherence, to augment and improve existing 2D trackers. Such geometric coherence cues are derived from epipolar geometry between images and other geometric invariances. We compare the performance of pre and post augmented 2D trackers on a variety of test sequences.

---

### **A camera-based human computer interaction system for users with disabilities (Demo and Poster)**

Wajeeh Akram, Laura Tiberii

Supervised By: Dr. Margrit Betke

Many people suffer from conditions that lead to deterioration of motor control making access to the computer using traditional input devices difficult. In particular, they may lose control of hand movement to the extent that the standard mouse cannot be used as a pointing device. Most current alternatives use markers or specialized hardware, for example, wearable devices, to track and translate a user's movement to pointer movement. These approaches may be perceived as intrusive. Camera-based assistive systems that use visual tracking of features on the user's body often require cumbersome manual adjustment. This work introduces an enhanced computer vision based strategy where features, for example on a user's face, viewed through an inexpensive USB camera, are tracked and translated to pointer movement. The main contributions of our work are (1) enhancing a video based interface with a mechanism for mapping feature movement to pointer movement that allows users to navigate to all areas of the screen even with very limited physical movement, (2) an evaluation of the camera mouse input paradigm as an input device and assessment of performance on elemental input tasks, and (3) providing a customizable, hierarchical navigation framework for human computer interaction (HCI). This framework provides effective use of the vision-based interface system for accessing multiple applications in an autonomous setting. Experiments with several users show the effectiveness of the mapping strategy and its usage within the application framework as a practical tool for desktop users with disabilities.

---

### **Event Prediction in a Hybrid Camera Network with Minimum Knowledge (Poster)**

U. Murat Erdem

Supervised By: Stan Sclaroff

One of the main advantages of an active camera, its large visual coverage area is also the source of a new problem: An active camera can only see a portion of its total visual coverage at any given time. In a significant number of cases the active camera is required to change its orientation to intersect its field of view by the object of interest. The new control aspect of the problem requires new decisions to be made for a foreseeable time horizon, i.e., "when" and "where" to foveate the camera in order to see an object. These are in addition to the already existing internal camera parameter decision problems like focal length and aperture. In a camera network the number of decisions increases dramatically by the number of cameras involved. Hence controlling the cameras such that events of interest are captured within parameters required by the visual task(s) is a

difficult challenge. This work presents a solution to this problem by proposing a prediction framework with the following contributions: First a predictive graph theoretical data structure is accrued over time which also has the ability to adapt to potential changes in the environment. At each camera data acquisition is performed at randomly visited preset PTZ (Pan-Tilt-Zoom) coordinates called "stations". When a trigger event at some station is followed by a target event at another station in a predefined time window, the data structure is updated to reflect the temporal associations of the detected events and their corresponding PTZ parameters. After significant amount of data is collected a novel SMC (Sequential Monte Carlo) based method starts to simultaneously predict camera parameters and track objects of interest in the system. The system works with minimal knowledge about the environment. No collective or individual calibration information is required. The stochastic event associations do not involve any appearance based cues sidestepping potential inter-camera color calibration pitfalls. Finally the evolving temporal event correlation structure also provides a basis to approximate the underlying temporal traffic patterns using graph theoretical approaches. In the current implementation the trigger and target events are respectively defined as blob and face detection in observing people in an office environment using a five PTZ camera network.

---

### **Vehicle gross classification** (Poster)

Ashwin Thangali, Vitaly Ablavsky

Supervised By: Stan Sclaroff

To track pedestrians accurately near cars we constrain the search for motion cues to specific image regions called receptive fields. The configuration and size of the receptive fields is determined by the vehicle type, orientation, and image size. In addition, receptive fields are clipped by opaque objects, such as other cars and therefore encode occlusion information. Vehicle class and an outline polygon are hence needed for each parked car to instantiate and compute occlusion relationships for its receptive fields. With small chip sizes (less than 100 x 100 pixels) and low dynamic range images, a sufficiently strong model is needed to achieve a reasonable performance. We choose to use template matching for vehicle classification. Exemplars are chosen for vehicle viewpoints and gross vehicle classes that frequently occur in our dataset, i.e., van, sedan, wagon, SUV, and truck. Matching is performed using edge images. We use normalized crosscorrelation to localize a vehicle template inside the test chip and chamfer distance to determine the best matching vehicle class template. The registered template also provides a vehicle outline polygon that is used to estimate inter-vehicle occlusions and eliminate the hidden parts of receptive fields.

---

### **Pedestrian-Vehicle Association in a Crowded Parking Lot** (Poster)

Vitaly Ablavsky and Ashwin Thangali

Supervised By: Stan Sclaroff

Our goal is to infer human activity in crowded settings such as a busy parking lot or interior of a supermarket. As a motivating example we would like to determine if a person X came from vehicle Y. Our source of information is a single fixed video camera. "Standard" activity recognition machinery does not work because large number of occlusions leads to severe tracking failures. We are developing a local occlusion-aware activity model and a method for global track inference. The local activity model is inspired by 3D but works purely in the 2D image plane. We solve global track inference and person-vehicle association as an instance of an assignment problem with domain-specific constraints.

---

**Stereoscopic Reconstruction of Foraging Bat Trajectories (Poster)**

Lisa Premerlani

Supervised By: Margrit Betke

To study the foraging behavior of bats, including the Brazilian free-tailed bat (*Tadarida brasiliensis*), we took infrared thermal videos of bats during warm weather nights in south-central Texas and used computer vision techniques to detect and track the bats. Our scene was bats foraging over a small body of water that was next to a corn field from which moths were emerging. Two cameras, approximately 20 meters apart, were used to take simultaneous videos of this scene for the purpose of doing a stereo reconstruction of the three dimensional (3D) trajectories of the bats. Our computer vision system implements an automatic, adaptive thresholding detection algorithm and a Kalman filter is used for tracking. To aid in the calibration of the cameras, we used heat producing calibration devices that we designed, constructed, and installed in the scene. Once the camera system was spatially calibrated and synchronized in time, we projected a ray from each camera center through the image of the bat in the 2D image plane and then we used triangulation to obtain the 3D coordinate of a foraging bat for each instance in time. Analysis of corresponding video frames recorded by the three cameras then produced the reconstructed 3D trajectory of the bat.

---

**Indexing Methods for Scalable Multiclass Recognition (Poster)**

Alexandra Stefan, Quan Yuan

Supervised By: Stan Sclaroff, Vassilis Athitsos

In many applications, such as face recognition and hand pose estimation, we need systems that can recognize a very large number of different classes. A common approach is to reduce multiclass recognition to the problem of constructing one classifier for each class; the task of this classifier is to decide if an object belongs to that class or not. Given an object to classify, all class-specific classifiers are applied and the object is assigned to the class of the classifier with the strongest response. A key problem with that approach is that recognition time is linear to the number of classes. We propose an indexing method for speeding up multiclass recognition. Our method is embedding-based: objects and classifiers are mapped to vectors in such a way that classifiers producing strong responses for an object tend to be mapped to vectors close to the embedding of that object. This way, a small number of candidate classifiers can be identified efficiently using well-known vector indexing methods.

---

**Networks****Amorphous Placement and Informed Diffusion for Efficient Field Monitoring by Autonomously Mobile Sensors (Poster)**

Hany Morcos

Supervised By: Azer Bestavros, Abraham Matta

Personal communication devices are increasingly equipped with sensors able to passively sample their surroundings. We envision a service that enables a community of users carrying such memory-limited devices to query the condition of various locations in the field in which they collectively roam. We show that existing techniques that rely on directed placement and retrieval (DPR), are viable approaches to implementing such a service, but only when the underlying network is well connected. Alternatively, we propose the use of amorphous placement and retrieval (APR), in which a cache management scheme is employed to store sensory samples locally, and an informed exchange of cached samples is used to diffuse the sensory data throughout the network, in such a

way that the answer to any query (targeting an arbitrary location in the field) is likely to be found close to the query origin. A salient characteristic in such a setting is the relationship between the probability of roaming a location in the field and the probability of querying that location. If roaming and query probability distributions do not match---which is the case in many settings---then an important determinant of the performance of APR is the manner with which cached field samples are collectively shared and managed. In that regard, we argue that knowledge of the distribution of query targets could be used effectively by an informed cache management policy to maximize the utility of collective storage of all devices. Using a simple analytical model, we show that the use of informed cache management is particularly important when the mobility model results in a non-uniform distribution of users over the field. We present results from extensive simulations which show that in sparsely-connected networks, APR is more cost-effective than DPR, that it provides extra resilience to node failure and packet losses, and that its use of informed cache management yields superior performance.

---

**netEmbed: A Framework for Solving Constrained Network Embedding Problems** (Demo and Poster)

Jorge Londoño

Supervised By: Azer Bestavros

Distributed applications are usually represented as overlay networks deployed on top of a hosting infrastructure. Traditionally such deployment occurs on a best-effort basis, but this is not enough when the performance of the application is constrained. Examples of such constraints are service level agreements or QoS specifications given on the links and/or the nodes of the overlay. This leads us to the problem of searching for feasible (and possibly optimal) embeddings of such overlays subject to a set of constraints, this is known as the "Network Embedding Problem". Some example applications where solving this problem is of interest include: - Mapping a network on a testbed - Deploying a distributed application on a GRID environment - Finding resources for deployment of an application on a Sensor Network In this work we present several techniques we have developed to find feasible solutions to the problem. As opposed to currently known techniques, ours does not sacrifice soundness, in other words, ours do not prune regions of space where feasible solutions may be found. We also present an evaluation of these techniques under a large number of synthetic scenarios where practical size topologies are considered.

---

**The Selfish Neighbor Selection Problem In Overlay Networks** (Demo and Poster)

Georgios Smaragdakis with Nikolaos Laoutaris (Harvard), John Byers (BU), Mema Roussopoulos (Harvard).

Supervised By: Azer Bestavros

We re-examine the problem of Overlay Network Creation, taking into consideration the existence of selfish overlay nodes. We develop a general Game-Theoretic framework that provides a unified approach to modeling Selfish Neighbor Selection (SNS) procedures, "wiring" for brevity, on behalf of such nodes. The model is general enough to take into consideration costs reflecting network latency and node preference profiles, the inherent directionality in overlay maintenance protocols and the limited recourses of each node in terms of bounded out- and in-degrees. Within this framework we formalize the notion of node's "best response" wiring strategy as a k-median problem on asymmetric distance, and use this formulation to obtain pure Nash equilibria. We show that selfish nodes can reap substantial performance benefits when connecting (following a "best response" wiring) to any overlay network, especially those composed of non-selfish nodes. Turning our attention to the performance of overlay networks that are dominated by selfish nodes, we experimentally examine the properties of such stable wirings on synthetic topologies, as well as on

real topologies and maps constructed from PlanetLab and AS-level Internet measurements. To capitalize on the substantial performance improvement of best response wirings for individual nodes and the emergent stable global wirings we design and deploy, EGOIST, an SNS-inspired prototype overlay routing network for PlanetLab. We show that best response is significantly more efficient than previous heuristic strategies, under multiple performance metrics, including delay and available bandwidth --- while at the same time achieving superior scalability and nearly as good performance as full-mesh based approaches. We also demonstrate how SNS yields multi-fold improvement of search performance compared to existing unstructured peer-to-peer file sharing systems. [For more information please refer to our research site: <http://csr.bu.edu/sns/>]

---

### **An Information Theoretic Approach to Anomaly Detection (Poster)**

Nahur Fonseca

Supervised By: Mark Crovella

The method of types is applied to sequences of packets extracted from links of aggregated source-destination flows of a large ISP. We show that this method needs to be correct due to the failure of the independence assumption at the level of packets and flows in the Internet. We then propose an empirical way to adapt the method of types to detect network anomalies in Internet packet traces.

---

### **The Cache Inference Problem and its Application to Content and Request Routing (Poster)**

Georgios Zervas

Supervised By: A. Bestavros, G. Kollios

In many networked applications, independent caching agents cooperate by servicing each others' miss streams, without revealing the operational details of the caching mechanisms they employ. Inference of such details could be instrumental for many other processes. For example, it could be used for optimized forwarding (or routing) of one's own miss stream (or content) to available proxy caches, or for making cache-aware resource management decisions. In this paper, we introduce the Cache Inference Problem (CIP) as that of inferring the characteristics of a caching agent, given the miss stream of that agent. While CIP is insolvable in its most general form, there are special cases of practical importance in which it is, including when the request stream follows an Independent Reference Model (IRM) with generalized power-law (GPL) demand distribution. To that end, we design two basic "litmus" tests that are able to detect LFU and LRU replacement policies, the effective size of the cache and of the object universe, and the skewness of the GPL demand for objects. Using extensive experiments under synthetic as well as real traces, we show that our methods infer such characteristics accurately and quite efficiently, and that they remain robust even when the IRM/GPL assumptions do not hold, and even when the underlying replacement policies are not "pure" LFU or LRU. We exemplify the value of our inference framework by considering example applications.

---

### **Generating Representative ISP Topologies From First-Principles (Poster)**

Chong Wang

Supervised By: John Byers

Understanding and modeling the factors that underlie the growth and evolution of network topologies are basic questions that impinge upon capacity planning, forecasting, and protocol research. Early topology generation work focused on generating network-wide connectivity maps, either at the AS-level or the router-level, typically with an eye towards reproducing abstract

properties of observed topologies. But recently, advocates of an alternative "first-principles" approach question the feasibility of realizing representative topologies with simple generative models that do not explicitly incorporate real-world constraints, such as the relative costs of router configurations, into the model. Our work synthesizes these two lines by designing a topology generation mechanism that incorporates first-principles constraints. Our goal is more modest than that of constructing an Internet-wide topology: we aim to generate representative topologies for single ISPs. However, our methods also go well beyond previous work, as we annotate these topologies with representative capacity and latency information. Taking only demand for network services over a given region as input, we propose a natural cost model for building and interconnecting PoPs and formulate the resulting optimization problem faced by an ISP. We devise hill-climbing heuristics for this problem and demonstrate that the solutions we obtain are quantitatively similar to those in measured router-level ISP topologies, with respect to both topological properties and fault-tolerance.

---

### **SNOW (SneakerNet Over Wireless) (Poster)**

Raymond Sweha

Supervised By: Azer Bestavros

In this project, we explore a routing protocol designed for mobile, ad-hoc networks that helps enhance requester/provider anonymity and plausible deniability. Unlike traditional P2P file sharing approaches that optimize for deliverability and global scalability, this work proposes a scheme that trades off deliverability for increased anonymity by exploiting the mobility of nodes. One salient characteristic that distinguishes this project from similar work in delay tolerant networks is that the protocol described herein opportunistically routes files without requiring scheduled communication or explicit addresses of the requester or provider, thus enabling some degree of anonymity and plausible deniability. We are motivated to explore this problem because such a protocol is a necessary component for enabling users to anonymously share sensitive files over ad-hoc networks formed by their mobile phones, thus effectively thwarting government censorship.

---

### **TCP Over CDMA2000 Networks: A Cross-Layer Measurement Study (Poster)**

Karim Mattar, Ashwin Sridharan (Sprint ATL) and Hui Zang (Sprint ATL)

Supervised By: Ibrahim Matta and Azer Bestavros

Modern cellular channels in 3G networks incorporate sophisticated power control and dynamic rate adaptation which can have a significant impact on adaptive transport layer protocols, such as TCP. Though there exists studies that have evaluated the performance of TCP over such networks, they are based solely on observations at the transport layer and hence have no visibility into the impact of lower layer dynamics, which are a key characteristic of these networks. In this work, we present a detailed characterization of TCP behavior based on cross-layer measurement of transport, as well as RF and MAC layer parameters. In particular, through a series of active TCP/UDP experiments and measurement of the relevant variables at all three layers, we characterize both, the wireless scheduler in a commercial CDMA2000 network and its impact on TCP dynamics. Somewhat surprisingly, our findings indicate that the wireless scheduler is mostly insensitive to channel quality and sector load over short timescales and is mainly affected by the transport layer data rate. Furthermore, we empirically demonstrate the impact of the wireless scheduler on various TCP parameters such as the round trip time, throughput and packet loss rate.

---

### **A Rule Based Decision Making Framework for Sensor Networks (Poster)**

Sowmya Manjanatha

Supervised By: Azer Bestavros

Sensor networks are emerging as popular means for real-time information gathering and processing in diverse contexts -- e.g., patient data monitoring in acute, chronic care and battlefield settings, environmental data from inhabitable terrains etc. We believe that for sensor networks to succeed in the envisioned heterogeneous environment, the process used to program and/or retarget them must be seamless. Towards that goal, we have built a general rule-based system that allows for unified sensory information processing and retrieval through the SnBench framework. We illustrate through a vital signs monitoring example that representation of sensor tasks as rules facilitates the ease of extension of the engine to new and emerging applications with specific needs.

---

### **An Energy-conscious Transport Protocol for Wireless Ad Hoc (Poster)**

Niky Riga

Supervised By: Ibrahim Matta

Within a recently developed ultra low-power ad hoc network system, we present a transport protocol whose goal is to reduce power consumption without compromising delivery requirements of applications. To meet its goal of energy efficiency, our transport protocol (1) contains mechanisms to balance end-to-end vs. local retransmissions; (2) minimizes acknowledgment traffic using receiver regulated rate-based flow control combined with selected acknowledgments and in-the-network caching of packets; and (3) aggressively seeks to avoid any congestion-based packet loss. Our extensive simulations demonstrate that our transport protocol meets its goal of preserving the energy efficiency of the underlying network. We also outline our current protocol implementation in Linux.

---

### **A Geometric Approach for Slot Alignment in Wireless Sensor Networks (Poster)**

Niky Riga

Supervised by: Ibrahim Matta, Azer Bestavros

Duty-cycling in Wireless Sensor Networks, led many researchers to design slotted protocols to run on top of these networks. In WSNs, like in any other distributed system, the clocks of nodes are not synchronized which puts a challenge in designing slotted based protocols. Most of the current approaches for aligning slots are based on running a synchronization protocol that gives guarantees on the maximum clock skew between nodes and then employ guard times in the beginning and ending of each slot to ensure proper alignment. This slot expansion is not appropriate for lightweight protocols like SDJS and DIP. In this paper we propose a new approach for slot alignment that gives a constant slot expansion independent of the maximum clock skew guaranteed by the synchronization protocol.

---

**Forwarding in Delay Tolerant Networks** (Poster)

Vijay Erramilli

Supervised By: Mark Crovella

We present some recent work on developing forwarding algorithms in delay tolerant networks. We describe in detail our measurement driven approach towards developing different algorithms, the motivation behind developing different algorithms and how they differ from one another in terms of performance on real data. External Collaborators: Augustin Chaintreau, Christophe Diot Thomson Research Paris

---

**Distributed Approach to Volume Anomaly Detection** (Poster)

Parminder Chhabra, Eric Kolaczyk and Clayton Scott.

Supervised By: Mark Crovella

In this work, we study whether the local view of traffic seen by an individual router (links connection to the router) is sufficient for accurate volume anomaly detection. Based on the characteristics of this problem, we infer normal traffic behavior via minimum measure sets. We show that the traffic observable at a router is not easily described in a parametric manner, such as by modeling with multivariate Gaussian distributions, but can be well characterized in a nonparametric way via minimum measure sets. Building on this observation we develop a simple distributed algorithm that allows individual routers to identify volume anomalies on their attached links. For sample data from a real network, we show that the performance of this method is on par with that of well known centralized approaches, without incurring the robustness and scalability drawbacks associated with centralized approaches.

---

**Operating Systems****Programming the Garcia Robot** (Demo and Poster)

Gerald Fry and Aaron Hughes

Supervised By: Richard West

Effort is now underway on a project to enhance extensibility and protection in embedded real-time system software using COTS components. This project involves porting the User-level Sandboxing mechanism to the Stargate XScale platform running Linux 2.6.11. Goals of this project include predictable and safe execution of real-time tasks associated with robotic sensors and actuators found on the Garcia robot manufactured by Acroname, Inc. Additionally, we seek to prove that widely available general purpose systems such as Linux can be extended to support real-time service in embedded platforms and that the programming interface for real-time task developers can be cleanly integrated with existing system services and API's.

---

**Wireless Network Intrusion Detection on the Sensor Network Workbench** (Demo and Poster)

Michael Ocean

Supervised By: Azer Bestavros

Wireless Network Intrusion Detection (WNID) is a well researched area and many tools exist to detect wireless anomalies. Network Intrusion Detection consists of two distinct phases (1) the monitoring and collection of Layer-2 or Layer-3 network packets and (2) the detection of anomalies via some analysis on the data collected. To actually derive value from WNID techniques, a third

component is required: an infrastructure to aggregate, collect and respond to detected events. Most tools provide some basic logging mechanism, but it is generally the role of the IT department to deploy these devices and generate some software (scripts) to aggregate and respond to events. Custom scripts/code to form a distributed wireless intrusion detection system tend to be device, package and deployment specific and thus cumbersome to maintain. A wireless network intrusion detection system is merely one specific instantiation of a (narrowly focused) sensor network. Rather than re-invent the wheel, we detail the inclusion of wireless network monitoring capabilities in our own Sensor Network workBench (snBench). The snBench provides an extensible framework for the rapid development and automated deployment of sensor network applications on a shared sensing infrastructure. The snBench's extensible architecture allows an engineer to quickly interface new modalities and sensing capabilities into the framework that may in turn be leveraged by novice SN programmers. This presentation gives an overview of the snBench and presents recent work to extend the snBench to integrate 802.11b/g wireless network monitors and a Pan-Tilt-Zoom camera network to develop a cross-modal network security monitoring service. Where as it is possible to custom build some of these functionalities without the use of the snBench, you wouldn't want to. Beyond the ease of development, deployment and infrastructure monitoring that snBench provides, the range of possible responses to a detected intrusion is greatly expanded when a cross-modal SN infrastructure is available. We also present current work on applying programming language verification techniques to snBench programs.

---

### **The Self-Organization of Component-based Systems for Dependable and Predictable Embedded Computing (Poster)**

Gabriel Parmer

Supervised By: Richard West

Failure in embedded systems is often disastrous, however the complexity of such embedded systems is steadily increasing making it impossible to prevent all errors pre-production. Hardware protection domains are often used to isolate different components of the system such that a failure in one does not necessarily effect others, increasing the robustness of the system. However, such protection domains incur a performance penalty on the system as a whole, and on specific applications. Many system structuring techniques conservatively place these protection domains statically, often without regard for application's predictability constraints. We propose a technique called Mutable Protection Domains (MPD) that dynamically places hardware isolation boundaries though-out the system to maximize useful isolation while still meeting application predictability and throughput constraints. In this way, failures in the system will be as contained as possible while still meeting application objectives. Specifically, in this poster we will discuss the algorithms and formalisms to decide when and where isolation should be present.

---

## **Programming Languages**

### **Quantifying over Class Relationships in Object-Oriented Programs (Poster)**

Joe Hallett

Supervised By: Assaf Kfoury

We present an object-oriented language that supports universal quantification over class relationships. This allows classes to be quantified beyond their ordinary type parameters with special, hidden, type variables. With hidden type variables, a single class can extend, and inherit methods from, infinitely many instantiations of another class. Moreover, with hidden type variables, we can express sophisticated class relationships such as covariance and contravariance, directly and with no additional language features. We can also elide redundant type parameters in type

instantiations, and define several other interesting and important class relationships. Instantiations of type parameters in our language are first-class types that are retained at run time and can be used in type-dependent operations. Although our formal analysis is directly applicable to the Fortress programming language, it is also relevant to the inclusion of hidden type variables in any type system with nominal subtyping and parametric polymorphism.

---

### **Verification of Network Flows Using a Type System With Constrained Polymorphism (Poster)**

Likai Liu

Supervised By: Azer Bestavros, Assaf Kfoury, Abraham Matta

We propose a typed domain-specific language TRAFFIC(X) for the specification and analysis of end-to-end network flow configurations. TRAFFIC(X) is inspired by HM(X), a framework that augments the familiar Hindley-Milner polymorphic type system for ML-like functional languages with a constraint system X. In TRAFFIC(X), the constraint system X can be instantiated to different constraint-rewriting rules, corresponding to the verification of different properties of network flows. To illustrate our methodology, we examine two examples: (1) proper nesting of tunneled streams by means of lossy or lossless compression and encryption; and (2) computing round-trip time using linear-programming constraints.

---

### **Safe Resource Sharing in Sequential and Concurrent Programming (Poster)**

Rui Shi

Supervised By: Hongwei Xi

The potential of linear logic in facilitating reasoning on resource usage has long been recognized. Recently, Zhu and Xi developed a type system that can effectively support safe explicitly memory manipulation through pointers. However, this linear type system does not properly support sharing of resources. As a consequence, it requires that resources be threaded through unctons, which can cause some serious difficulties in practice and thus severely restrict the use of linear types. We propose to develop a general type-theoretical framework for supporting safe sharing of resources in practical programming. For this purpose, we introduce and then formalize a modality, which we call sharing modality. A naive treatment of this sharing modality leads to unsoundness due to the issue of non-reentrancy. To address this problem, we are to employ a notion of types with effects to reflect at the level of types whether a function is reentrant or non-reentrant. An application domain we are particularly interested in is OS implementation and extension, where concurrency is ubiquitous. Thus, we are highly motivated to develop (advanced) type theory that can facilitate the construction of safe and reliable concurrent programs. A major goal we want to achieve is to convincingly substantiate the claim that linear and dependent types can be employed effectively in practice to support safe manipulation of (linear) resources shared by multiple threads.

---

### **A Theorem Proving Language based on constraint types (Poster)**

Kevin Donnelly

Supervised By: Hongwei Xi

We present a core theorem-proving language based on indexed types with impredicative quantification and constraint types and show that this language forms a consistent logic. We argue that this core language can be the basis for a language that is more user-friendly than previous theorem-proving languages based on dependent types and that can be easily embedded into a programming language for use as a program logic.

---

## Security

### **Conditional Computational Entropy** (Poster)

Chun-Yuan Hsiao

Supervised By: Leonid Reyzin

We study conditional computational entropy: the amount of randomness a distribution appears to have to a computationally bounded observer who is given some correlated information. By considering conditional versions of HILL entropy (based on indistinguishability from truly random distributions) and Yao entropy (based on incompressibility), we obtain: 1) a separation between HILL entropy and Yao entropy in the shared random string model (improving on Wee's 2004 separation in the random oracle model); 2) the first demonstration of a distribution from which extraction techniques based on Yao entropy produce more pseudorandom bits than appears possible by the traditional HILL-entropy-based techniques; 3) a new, natural notion of unpredictability entropy, which implies conditional Yao entropy and thus allows for known extraction and hard-core bit results to be stated and used more generally.

---

### **Intrusion-Resilient Secure Channels** (Poster)

Scott Russell and Robert McNeerney Jr.

Supervised By: Gene Itkis

Alice in Alaska and her friend Bob in Boston can communicate confidentially using a secure channel that employs some encryption algorithms and agreed-upon secret keys. Unfortunately, a malicious party Eve may somehow obtain either or both of their secret keys. If this occurs, then from that point onward Eve will be able to decipher their supposedly private conversation. We propose a new primitive called an Intrusion-Resilient Secure Channel that automatically restores the confidentiality of their messages without repeating the potentially cumbersome initial key agreement. In addition to a formally defining this new primitive, we demonstrate a simple, generic construction using ordinary, existing public key encryption algorithms. We prove a bound on the security of our construction under the assumption that Eve is able to observe, but not interfere with the communication between Alice and Bob. We also describe how intrusion-resilient secure channels can be used to improve the security of other protocols.

---

### **Private Range Queries** (Poster)

Nenad Dedic and Scott Russell

Supervised By: Leo Reyzin

The ability to easily search public and proprietary databases via the Internet is beneficial for businesses and individuals. For competitive, personal, or other reasons, searchers may prefer to keep their queries and results confidential, even from party hosting the database. We demonstrate a protocol for privately executing simple range queries of the form "What values in the table are between 100 and 800?" against a server-hosted numerical table. To the best of our knowledge, this is the first such range query protocol with provable privacy guarantees for both the querier and the server. Querier privacy means the server learns nothing about the query, including the answer to it. Server privacy means the querier learns the correct answer to the query but no additional information about the table. Our protocol for honest-but-curious querier and server combines existing cryptographic primitives for private comparison and private retrieval by address with the classic binary search tree data structure. To handle malicious users we provide a novel version of universal hashing that is computable under encryption.

---

**Robust Fuzzy Extractors (Poster)**

Bhavana Kanukurthi

Supervised By: Leonid Reyzin

In cryptography, we often assume the existence of "uniformly distributed secret keys". An interesting question, though, is how to generate, store and retrieve such strings. Sources of strings that are neither "highly random" nor "reliably retrievable" are plenty. An example of such a source would be a person's iris scan. Surprisingly, such sources can be transformed into sources that can be used in cryptographic applications. Robust fuzzy extractors, which enable two parties who share a correlated random variable to exchange a uniformly distributed secret, mathematically model this goal. We present our construction which improves previous results in the length of the uniformly distributed secret generated.

---

**On Trading Secret Random Bits for Public Ones (Poster)**

Nenad Dedic

Supervised By: Leonid Reyzin, Danny Harnik (Technion)

We study the question of when a one-way function can be used with fewer input bits than prescribed. We show that by using *public* randomness it is possible to use a *regular* one-way function (i.e. a function in which every output has the same number of preimages) with only as much secret randomness as the entropy in the function's the output. Moreover, we show that the use of public randomness is essential as the total amount of randomness required for using a one-way function cannot be reduced using black-box techniques (even for regular functions). We generalize this result to hold for any function where all of its outputs are "heavy" (have a large number of preimages).

---

**Theory****Games on the Sperner Triangle (Poster)**

Kyle Burke

Supervised By: Shang-Hua Teng

We create a new two-player game on the Sperner Triangle based on Sperner's lemma. Our game has simple rules and several desirable properties. First, the game is always certain to have a winner. Second, like many other interesting games such as Hex and Geography, we prove that deciding whether one can win our game is a PSPACE-complete problem. Third, there is an elegant balance in the game such that neither the first nor the second player always has a decisive advantage. A playable applet for the game is available at <http://cs-people.bu.edu/paithan/spernerGame/>

---

## **Unique Games and Approximation Algorithms (Poster)**

David Charlton

Supervised By: Peter Gacs

The  $P \neq NP$  hypothesis lets us classify many algorithmic problems as "intractable." The PCP (Probabilistically Checkable Proof) Theorem builds on top of this, showing that many NP-Complete problems are hard not just to solve exactly, but to even approximate. We investigate Khot's "Unique Games Conjecture," an as-yet unproven strengthening of the PCP Theorem that greatly increases the approximation gap for many problems beyond what was previously known. Our study particularly focuses on the proof techniques and types of reductions that would be needed to prove the conjecture.

---

## **Hierarchies in PSPACE (Poster)**

Ben Hescott

Supervised By: Steve Homer

We review the jump hierarchies within PSPACE. Typically a jump hierarchy in PSPACE is constructed level by level by adding more resources (nondeterminism); this is believed to add the ability to solve more problems. The delta levels of the polynomial hierarchy are created just this way by starting with P and repeatedly adding a SAT oracle at each level. We create a different hierarchy --- one built top down. Under a reasonable assumption (the NP-machine hypothesis) we create the reverse polynomial hierarchy. We start with PSPACE complete sets and create jump hierarchies that extend downward. This hierarchy can be thought of as a set of incomplete hierarchies, at the bottom level the set cannot be complete for any level of the original polynomial hierarchy until enough nondeterminism is added to reach all of PSPACE.