



Message from the Chairs



**Kathy McKeown,
Professor &
recent Chair**

I have had great pleasure serving as department chair from July 1997 through December 2003, despite the difficulties that came with it. As I look back over the last six years, I am particularly pleased with a number of changes in the Department. The faculty has grown from 20 to 32, establishing critical mass in several key areas. The faculty now meets once a year on a retreat that focuses on long-term issues for the Department. Many strong student organizations have flourished, such as

the ACM Student Chapter, the Research Liaison Program, the student run TA group, and the Women in Computer Science group. A new Center for Computational Learning Systems was established and will augment ongoing research in the CS Department. Most importantly, the feeling of community spirit in the Department has created an environment for positive change.

As of January 1, 2004, Henning Schulzrinne will take over as department chair. He and I have worked closely over the last semester and it is very clear that he has the energy, vision, and direction to lead the Department to even better places.

As I look to the future, I am looking forward to spending more time with my students and on my research... and to receiving less email :) Best wishes Henning!



**Henning Schulzrinne,
Professor
& Chair**

Welcome to the second edition of the newsletter for the Department of Computer Science at Columbia University. The Department has made great strides in the past few years, allowing us to build on the solid

foundation laid by Professor McKeown and the prior chairs. We are fortunate to be able to welcome Mihalis Yannakakis as the new Percy K. and Vida L. W. Hudson Professor of Computer Science. He will significantly strengthen our efforts in the theory of computing (*see article on pg. 3*). The Master's program, recently thoroughly modernized, has the largest group of students in the Department's history, with a total of 398 students enrolled.

The Computer Science community at Columbia was deeply saddened by the unexpected death of one of our faculty, Professor Andrew Kosoresow, in June (*see pg. 7*).
(continued on page 6)

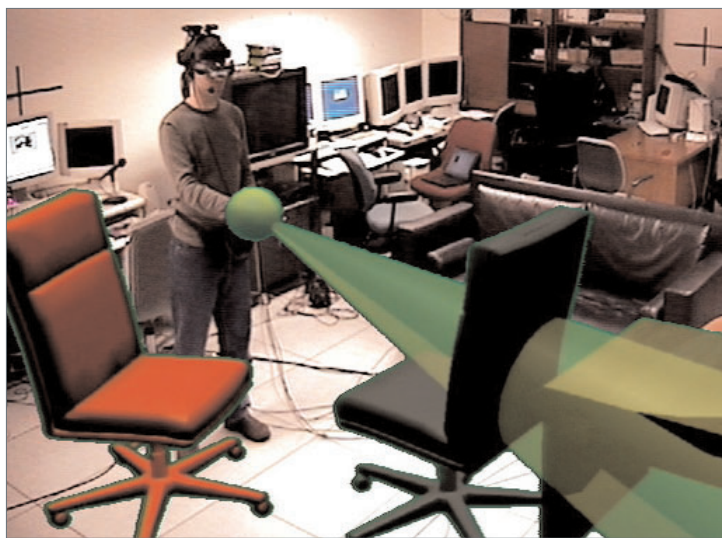
Computer Graphics & User Interfaces Lab



Steve Feiner

Professor Steve Feiner's Computer Graphics and User Interfaces Lab is involved in wide-ranging and often interdisciplinary research. **One of the most equipment-intensive labs in CS, they focus on interactive graphics – with an emphasis on virtual and augmented reality, mobile and wearable computing, and knowledge-based design of graphics and multimedia.** Here we provide brief overviews of just two of their current projects to offer a sense of the scope of their research.

<http://www1.cs.columbia.edu/graphics>



SenseShapes seen through a position- and orientation-tracked display that overlays graphics on the user's view of the world.

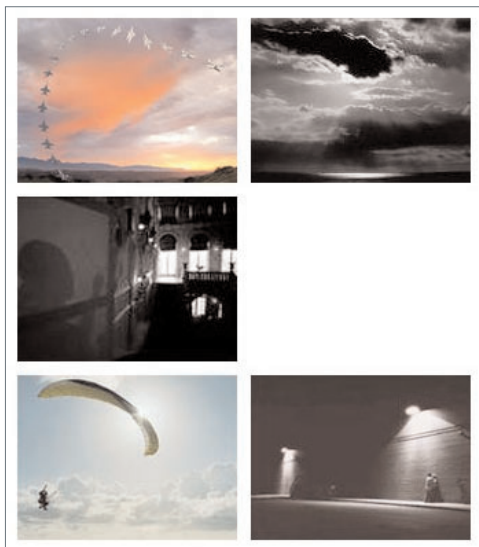
SenseShapes

Two lab members, Staff associate Alex Olwal and Ph.D. student Hrvoje Benko have been developing a set of statistical geometric tools designed to identify the objects being manipulated in an augmented reality system. In augmented reality, virtual objects are overlaid on the user's view of the real world. This work helps make it possible for users to refer to real and virtual objects multimodally, through combinations of speech and gesture. SenseShapes are volumetric regions of interest that can be attached to tracked parts of the user's body to provide

valuable information about the user's interaction with objects. To assist in object selection, SenseShapes generate a rich set of statistical data about objects and their spatiotemporal relationships to each volume, and the system dynamically chooses which data to consider, based on the current situation.

<http://www.cs.columbia.edu/graphics/publications/senseshapes.pdf>

Images laid out on a regular grid, using an automated layout system based on WeightMaps.



WeightMaps

Another lab member, Ph.D. student Simon Lok, working with M.S. student Gary Ngai, is exploring ways to automate graphical layout, the process of determining the size and position of the visual objects in an information presentation. They have developed the WeightMap, a bitmap representation of the visual weight of a presentation, along with algorithms that use WeightMaps to allow an automated layout system to evaluate the effectiveness of its layouts. This approach is based on the concepts of visu-

al weight and visual balance, which are fundamental to the visual arts. Each object in a layout has a visual weight, and a WeightMap is created that encodes the visual weight of the layout. The researchers use image-processing techniques, including pyramids and edge detection, to efficiently analyze the WeightMap for balance, and derivatives of the sums of the rows and columns to generate suggestions for improving the layout.

<http://www.cs.columbia.edu/graphics/publications/p101-lok.pdf>

The Network Computing Laboratory



Jason Nieh

The Network Computing Laboratory's (NCL) research interests have a strong

experimental focus, and are generally in building software systems that demonstrate significant functional and performance improvements with real applications.

The lab is directed by Professor Jason Nieh and includes a team of eleven PhD students and more than twenty Masters and Undergraduate students. Current research spans a broad range of areas, including operating systems, mobile computing, thin-client computing, web and multimedia systems, and performance evaluation.

One current NCL research project is Zap, which is a system developed in the NCL that enables unmodified Linux applications to be transparently suspended on one machine and migrated and resumed on another machine. Zap provides functionality analogous to the suspend/resume feature of a laptop computer, except that Zap can suspend on one machine and resume on a completely different machine.

Furthermore, Zap provides this functionality at a much finer-granularity, enabling even a single application to move from one machine to another. The tremendous potential benefits of Zap include fault resilience by migrating processes off of faulty hosts, better system response time by migrating processes closer to their users, dynamic load balancing by migrating processes to less loaded hosts. Other benefits include improved service availability and administration by migrating processes before host maintenance so that applications can continue to run with minimal downtime.

Another recent NCL research project is Certes (CliEnt Response Time Estimated by the Server), which is an online server-based mechanism that allows web servers to estimate client perceived response time, as if measured

at the client web browser. Certes provides accurate response time measurements without any modifications to HTTP servers, web pages, or client web browsers. It does not require any probing or third party sampling infrastructure. It works for any web content, not just HTML. Furthermore, Certes runs online in constant time with very low overhead. Because it can be used effectively at web sites and server farms to verify compliance with service level objectives, Certes is being deployed by IBM on every IBM-supported operating system.

<http://ncl.cs.columbia.edu>

The Languages & Compilers Group



Al Aho



Stephen Edwards

The Languages and Compilers group, headed by Professors Al Aho and Stephen Edwards, aims to develop tools that enable designers to quickly and correctly design complex hardware and software systems.

The central idea is to raise the level of abstraction presented to programmers to allow them to better cope with complexity and get it right. This is achieved by augmenting existing languages and compilers and by creating new domain-specific programming languages. They currently have six PhD students working on these efforts.

One research thrust is to attack the complexity in embedded system design. Embedded systems are unconventional computer systems designed for specific applications, such as automotive control or cellular phones. They are challenging because of tight constraints on performance, power, size, and weight, and their unconventional mix of hardware and software. Ongoing projects in this area include a domain-specific language for device drivers and

hardware and software generation from the Esterel real-time language. This work has already produced the Columbia Esterel Compiler, the first open-source implementation for the language.

Another research thrust aims to improve the reliability of software systems by automating techniques for validating programs. Projects in this space include developing easy-to-use tools that programmers can use to impose constraints on their code, an automatic mechanism for verifying that the code generated by a compiler correctly implements the program it was given, and techniques for verifying that a program has not been hijacked by an invader.

The final research thrust focuses on the emerging area of quantum computation. This is an exciting area truly in its infancy: quantum computers exist, but are almost comically

primitive. A handful of quantum algorithms are known, but new ones are very difficult to develop. Their focus is on creating languages and compiling techniques for implementing quantum algorithms on emerging quantum computing technologies.

<http://landc.cs.columbia.edu>

CLASS

The Columbia Center for Computational Learning Systems



Vladimir Vapnik
Professor



Yoav Freund
Senior Research Scientist



Phil Long
Senior Research Scientist



John Ioannidis
Senior Research Scientist



Vasileios Hatzivassiloglou
Research Scientist



Christina Leslie
Research Scientist

In May 2003 an ambitious new research center, the Columbia Center for **Computational Learning Systems** (CLASS), was launched with funding from the Computer Science Department, The Fu Foundation School of Engineering and Applied Science, and the Columbia central administration. CLASS will be leveraging Columbia's current strengths in the areas of data mining, natural language processing, learning, and network security to extend the effective size and scope of the CS Department's research and interdisciplinary efforts. Their stated goal – "to be a world leader in learning and data mining research and the application of this research to a wide variety of areas including bioinformatics, systems security and natural language understanding – as well as other emerging areas" – is already well underway.

What makes the center's approach so compelling? Using high-order mathematics, CLASS researchers are working to create methods optimized to handle complex multi-dimensional sys-

tems. Problems being tackled by CLASS can be very large – over 100,000 input dimensions – but novel methods can handle them with relatively little set up effort. Examples of this type of system are financial markets, power grid control or drug discovery. This potential for analysis and control of naturally occurring complex systems is considered by many to be a conceptual breakthrough. The Center has quickly become involved with many other departments at Columbia and elsewhere. New projects and proposals have already commenced with participants from around the world. Besides its natural interaction with so many diverse disciplines at Columbia, CLASS also has the potential to develop applications that could truly impact our lives.

A primary asset of the Center is its remarkable line-up of researchers. CLASS's Director, **David Waltz**, former President of the NEC Research Institute, and a leader in the field of artificial intelligence, is as personable as he is brilliant. He has moved quickly to hire a number of world-class researchers in

The prospects for CLASS are extraordinary. Besides its natural interaction with so many diverse disciplines at Columbia – and elsewhere – and the potential applications that could truly impact our lives, it might very well change the way we think about science itself.

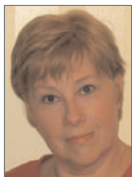
the fields of machine learning and data mining, most notably **Vladimir Vapnik** – renowned for his seminal work in the area of support vector machines, **Yoav Freund**, a leading machine learning researcher, renowned for his "Boosting" methods, **John Ioannidis**, a leader in systems security research, and **Phil Long** from the Genome Institute of Singapore. **Christina Leslie**, from the CS Department, recently accepted an offer for a full-time Research Scientist position with CLASS, and will be working in the Computational Biology Group with Freund and Long. **Vasileios Hatzivassiloglou**, who had been a Research Scientist in the CS NLP group, has joined CLASS to continue and expand his research on natural language and Web data processing, and **Martin Jansche**, a Post-Doc from Ohio State, has joined to collaborate with Professor Julia Hirschberg on speech learning. The Center is not only expanding its team – it also is expanding its facilities. Besides the suite of offices

that they are now occupying on the 5th floor of Mudd and joint-research labs in the CS and CEPSPR Buildings, they will be further expanding into an additional 4,200 square feet of office space currently being configured for them in the Interchurch Building on 120th and Claremont.

At the end of January 2004, CLASS hosted a two-day site visit here at Columbia to present its proposal as a finalist in the National Science Foundation Centers for the Science of Learning program. The proposal is for the creation of a "Center for Advanced Empirical Inference Studies," a huge project with over 100 different pieces. (If funded, this center would absorb the current CLASS center and would become the main center.) Computer Science faculty also involved with the proposal are Kathy McKeown, Director of the NLP Group (who was instrumental last year in helping establish the center while chair of the CS Department), Luis Gravano, from the Database Group, Betsy Sklar from the Autonomous Agents Group, and Owen



Martin Jansche
Post Doc



Nancy Burroughs Evans
Departmental Administrator

Rambow, Research Scientist with the NLP Group. Going well beyond the mathematical core of the proposal, there are also parts on software tools, education – with involvement of Sklar, Jack McGourty, associate Dean at the Fu Foundation, Susan Lowe from Teacher’s College, and others. Additional components focus on biology, philosophy and more. Ten other institutions besides Columbia are also partnering in the project, including Princeton University, NYU, NEC Labs, University of Maryland, University of Minnesota and GlaxoSmithKline. International partners include the Royal Holloway University of England, Max Planck Institute (in Germany), National Taiwan University, Riken Brain Institute of Japan, and the Institute of Russian Language. The judgment on this funding should be received in March.

<http://www.seas.columbia.edu/class>

Q & A



David Waltz
Director

To better understand the potential of this research and to find out why this has been generating so much excitement, we recently sat down for a talk with David Waltz, Director of CLASS to discuss the Center.

The following comments relate to the pending NSF proposal, based on Vladimir Vapnik’s high-dimensional learning methods.

Q: *I recall hearing a hypothesis that this approach has some parallels to the workings of the human mind – could you explain?*

A: This method is the antithesis of the ordinary western method of “divide and conquer” for solving problems – where you essentially control the dimensionality of a problem by picking the most important aspect first to guide your solution and once you’ve solved this part, looking for the remnant error – what problems haven’t been covered – and then taking a whack at that with the next most important dimension, repeating until you eventually cover everything.

But in the real world, for many engineering design problems – or also for ethics or legal problems – you can’t just choose one dimension to guide the whole thing. You sometimes have to balance and treat very seriously many dimensions and we believe this can be solved by some analogue in the human mind.

Q: *From a philosophical perspective, what is so different about the way CLASS is approaching these complex issues?*

A: In many ways traditional science has taken the view of “Occam’s Razor”; that is, you want to explain things by the simplest expression with the smallest number of variables. There also is a corollary: people have come to believe because Occam’s Razor has proved so successful in explaining the physical world, *everything* in the world should be describable by very simple formulas, and so they look only at explanations with small numbers of terms, small degree polynomials, and simple differential equations.

But our view, expressed in this proposal, is that this is like looking under the streetlamp for your lost keys, because that’s where the light is. What we’re trying to do is to attack the many, many different types of problems where that model is completely inapplicable – for example biological systems, financial systems, social systems in general, organizational systems. These problems are inherently high-dimensional and are not going to be described by a simple equation, so we really need a new kind of philosophy of science to understand them.

Rocco Servedio has been awarded a Faculty Early Career Development Award (CAREER) from the NSF, in the Theory of Computing Program. The title of Servedio's project is "Efficient Learning Algorithms for Rich Function Classes".

Jonathan Gross' most recent book "Handbook of Graph Theory" was just published by CRC Press.

Jason Nieh was awarded a Department of Energy Early Career Principal Investigator award in Applied Mathematics, Collaboratory Research, Computer Science, and High-Performance Networks. His project is titled "Migration Mechanisms for Large-Scale Parallel Applications".

Edward Shortliffe has been appointed to the oversight committee for the Division of Engineering and Physical Sciences (DEPS), National Research Council, National Academy of Sciences. Boards in DEPS include the Computer Science and Telecommunications Board, on which Shortliffe served from 1990-1996.

Steve Feiner received an IBM Faculty Award for his research on wearable computing for healthcare providers.

Tony Jebara has been awarded a Faculty Early Career Development Award (CAREER) from the NSF. His project is titled "Discriminative and Generative Machine Learning with Applications in Tracking and Gesture Recognition".

Tal Malkin has also been awarded a Faculty Early Career Development Award (CAREER) from the NSF, in the Theory of Computing Program. The title of Malkin's project is "Strengthening Cryptography by Reducing Assumptions about the Adversary".

Cisco Systems and Nortel Networks donated a new IP-based telephone system to the department. The system uses technology developed at **Columbia University**.

Tony Jebara was recently awarded a research grant from the NSF under its ITR program. The title of the project is "Representation Learning: Transformations and Kernels for Collections of Tuples".

Yoav Freund has won the 2003 Godel award in theoretical computer science. This was a joint award with Rob Schapire of Princeton University.

Tony Jebara's book "Machine Learning: Discriminative and Generative" was recently published by Kluwer.

Ravi Ramamoorthi received two grants to support his research on real-time rendering, from the NSF (Real-Time Rendering and Visualization of Complex Scenes) and Intel Corporation (Real-Time Rendering and Interaction with Complex Illumination and Materials).

"A Kernel between sets of Vectors" by **Risi Kondor** and **Tony Jebara** won best paper at the International Conference on Machine Learning, August 2003.

The technology developed by the Columbia IDS project, led by **Sal Stolfo**, has been licensed for commercial development by Columbia University to System Detection, Inc.

Henning Schulzrinne received the Mayor's Award for Excellence in Science and Technology from New York City Mayor Michael Bloomberg.

Andrew Miller and **Peter Allen**, along with co-authors **Steffen Knoop** and **Henrik Christensen**, won the Wegbreit Best Manipulation Paper Award at the 2003 IEEE International Conference on Robotics and Automation for the paper "Automatic Grasp Planning Using Shape Primitives".

Al Aho, was honored by The Society of Columbia Graduates (SOCG) with its Great Teacher

Award. Aho was recently named to the Department's newly created position of Vice-Chair for Undergraduate Education.

Kathy McKeown has been recognized as an ACM Fellow. This recognition is for having made significant advances in theoretical as well as practical computing that are having lasting effects on the lives of citizens throughout the world.

Angelos Keromytis was awarded a grant of 3 IXP 1200 network processors from Intel and a grant from DARPA to refurbish his SOS lab. He also has been awarded two grants from Cisco and also another from Intel.

Zvi Galil has been elected to the National Academy of Engineering. Each year the NAE salutes leaders in engineering for their lifetime dedication to their field and their commitment to advancing the human condition through great engineering achievement and/or through innovation in engineering and technology education.

Peter Belhumeur, **Steve Feiner**, and **Ravi Ramamoorthi**, along with colleagues at the University of Maryland and the Smithsonian Institution, received a NSF ITR Award for "An Electronic Field Guide: Plant Exploration and Discovery in the 21st Century".

(Message from the Chairs continued from page 1)

During the new year, and the beginning of my term as department chair, I look forward to working with the other faculty and the departmental staff on a number of projects that will continue the upward trajectory of the Department. We are revising the undergraduate program for the Computer Science major. The goal will be to attract the best students at Columbia and elsewhere to the Department. We want to provide a broad foundation both for students that are entering the workforce with a Bachelor's degree as

well as those continuing on to graduate study. The program requirements will be streamlined, to simplify advising and give more choices to students. This is becoming more important as computer science becomes a core component of anything from the financial industry to making movies and music. Many more students will be working in integrating and tailoring systems to support new uses, rather than the traditional notion of building programs from scratch.

We will continue to try to improve how we tell alumni, friends and colleagues about research and education within

the Department. We hope to produce this newsletter twice a year. To make CS@CU as relevant as possible, I would appreciate hearing from readers about features they'd like to see. If you are interested in receiving more timely updates on departmental news, you can subscribe to the mailing list at: <http://lists.cs.columbia.edu/mailman/listinfo/cucs-news>.

Our new research computing facilities director, Daisy Nguyen, will continue her work to modernize our computing facilities and phone system. New video collaboration facilities will shortly be available in our meeting rooms and lounge, allowing us

to work more efficiently with distant and close-by colleagues.

We are slowly adding functions to our new web-based departmental information management system. In particular, during the upcoming year, alumni will be able to set up email forwarding and more readily get in touch with other CUCS alumni.

This fall, we will be celebrating the 25th anniversary of the Department, involving current members, alumni and guests. Please look for additional information in the months to come.

Send comments, suggestions & news items to: chair@cs.columbia.edu

Super Theory Day

On Friday, May 14th Columbia University will be hosting a "Super Theory Day."

Presented in conjunction with NYU and IBM Research, it will be held in the Altschul Auditorium in the School of International and Public Affairs (SIPA) located at 117th Street and Amsterdam Avenue.

The event is a special edition of the semi-annual conference,

aimed at bringing together people in the New York Metropolitan area for one day of interaction and discussion. The Theory Day features several (usually 4-5) hour-long presentations by leading theoretical computer scientists about state-of-the-art advances in various areas. Some presentations give a survey of the latest advances in some area, while others may concentrate on a particular result.

Zvi Galil, Dean of Columbia's Fu Foundation School of Engineering and Applied Science, will be delivering the

opening remarks followed by speakers and panelists from around the nation. Scheduled to participate are Shafi Goldwasser, from MIT and the Weizmann Institute of Science in Israel, Richard Karp from UC Berkeley, Prabhakar Raghavan of Verity and Stanford University, Peter Shor of MIT, and Avi Wigderson from the School of Mathematics, Institute for Advanced Study at Princeton University.

Mihalis Yannakakis, who recently joined the Columbia CS Department will be moderating

a panel discussion on "The Future of CS Theory".

The meeting is free and open to everyone, particularly students, who are strongly encouraged to attend. Lunch and a T-shirt will be provided to those who RSVP.

Additional details and contact information for registering are located at: <http://www.cs.columbia.edu/theory/sp04.htm>

New Faculty



Mihalis Yannakakis

He joined the Columbia Computer Science Department in January 2004 as the Percy K. and Vida L.W. Hudson Professor of Computer Science.

Considered one of several leading theoretical computer scientists of his generation, he has made major contributions to

both algorithms and to computational complexity in a variety of areas, including database theory, combinatorial optimization, complexity theory, theory of approximability and verification. He was Director of the Computing Principles Research Department Laboratories at

Bell Labs, where he worked for more than for 20 years after receiving his Ph.D. in CS from Princeton. He briefly joined Avaya Labs and then spent one year as a CS Professor at Stanford University, before coming to Columbia.

<http://www.cs.columbia.edu/~mihalis>

In Memoriam



Andrew P. Kosoresow

Assistant Professor of Computer Science, and a beloved member of the Columbia University Computer Science Department, died on June 1st, 2003, at the age of 39. He died of heart failure while at home here in NYC.

A native New Yorker, Andrew received his BA in Computer Science in 1985 from Columbia

before leaving for the West Coast to attend Stanford University, where he received his Ph.D. in Computer Science in 1996. He taught at Stanford and then at the University of New Mexico before coming back to Columbia in 1997, where he stayed until his death.

A wonderful teacher, he was a great inspiration to the many students he touched during his six years as instructor in the Department. He had the unique ability to promote an atmosphere of belonging and openness, which extended well beyond the classroom. He acted as advisor to many undergraduates, regularly taking time, often well after office hours, to personally make sure that they understood and felt confident about their academic options, as well as helping them cope with the pressures of adjusting to the rigors of the CS curriculum.

He single-handedly revived the Columbia ACM student chapter, and through constant attention

as their faculty advisor, brought them to national recognition when they won the Excellence Award for Outstanding Chapter Activities in 2001 and the Best Community Service Award in April, 2003. He also had a great influence on the Department in his role as "TA czar", where he reorganized the system for assigning TAs to the myriad of classes being taught each semester. Andrew established, and was the teacher for the course, Computer Science Education, which taught TAs how to better teach CS. In recognition of his teaching skills and commitment to students, he received the 2001 Kim Award for Faculty Involvement from The Fu Foundation School of Engineering & Applied Science.

Andrew's research interests were in the area of Artificial Intelligence, including agent-based systems, distributed artificial intelligence, emergent computation, and machine learning; analysis of algorithms

including distributed and parallel algorithms, online algorithms, and randomized algorithms; and computer security including intrusion detection and cryptography. Applications of interest included epidemiology, economics, ecology, and archeology.

The Andrew P. Kosoresow Memorial Fund has been established by the Columbia CS Department. A new award, The Andrew P. Kosoresow Memorial Award for Outstanding Performance in TA-ing and Service, will be given annually.

Those who wish to give in Andrew's name can send a check to the Department. Please make the check payable to: Columbia University Department of Computer Science Attention: Mary Van Starrex 450 CSB, Mail Code 0401 New York, NY 10027

Further recollections of Andrew can be found on a new Columbia CS webpage dedicated to his memory: <http://www.cs.columbia.edu/apkmemorial.html>

The Network Security Lab



Angelos Keromytis

The Network Security Lab (NSL) is directed by Professor Angelos Keromytis, with participation from Vishal Misra, Tal

Malkin, Dan Rubenstein, Moti Yung, and John Ioannidis. Their research is focused on system and network security and survivability, with other research in efficient cryptographic mechanisms and protocols, access control policies, and intrusion detection.

One current project, the Worm Vaccine, addresses the problem of malicious worms and viruses. These have been particularly prevalent on the Internet in the past few years, causing significant network

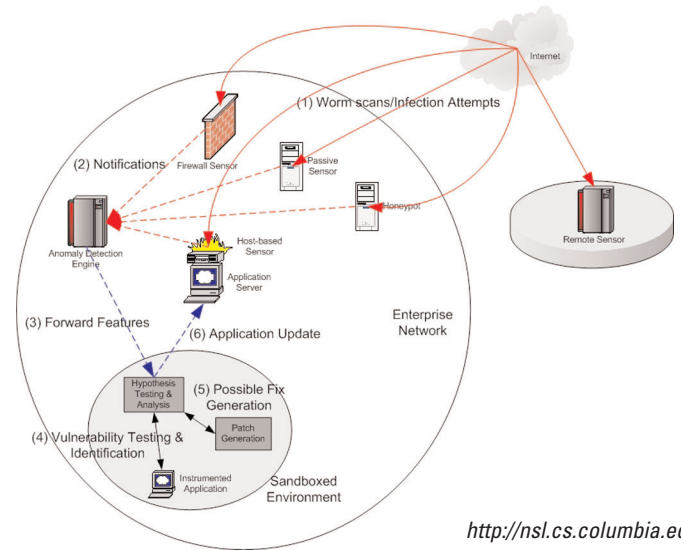
outages and monetary damage (reported to be in the billions of dollars for each major virus). Other researchers focus on filtering such attacks in the routers, in a manner akin to the way anti-virus software works on PCs. The work in NSL focuses on automatically identifying and fixing the software vulnerabilities that these viruses exploit.

The Worm Vaccine system uses a combination of "honeypots" (machines configured to look like legitimate servers on the network, whose only purpose is to attract attackers and monitor their activities), "sandboxing" (a mechanism for catching errors or abnormal behavior on a piece of software), and source-code rewriting.

When a virus first appears on the network and tries to exploit a software flaw to infect new computers, the Worm Vaccine not only detects this, but automatically generates and tests a

set of software patches that close the hole. As a result, the system remains secure and available (i.e., does not crash) during a virus outbreak. An organization can deploy this system on their own network, without coordinating with others, and reap the benefits of increased protection.

Currently, the students at NSL are investigating additional mechanisms that will allow the Worm Vaccine to evolve into a system that can detect and fix other types of software problems, as well as develop a theory of "correct process termination" in the presence of faulty code.



<http://nsl.cs.columbia.edu>

Department of Computer Science
Columbia University
1214 Amsterdam Avenue
Mailcode: 0401
New York, NY 10027-7003

ADDRESS SERVICE REQUESTED

NON-PROFIT ORG.
U.S. POSTAGE
PAID
NEW YORK, NY
PERMIT 4332