September, 2003

CRA Workshop on Research Related to National Security: Report and Recommendations

by Kathleen McKeown, Lori Clarke, and John Stankovic (Organizers)

Introduction

The Computing Research Association (CRA) hosted a workshop in September 2002 to develop recommendations that will strengthen the research infrastructure in areas of critical importance to national security. The workshop was supported by the National Science Foundation.

The workshop focused on three general topics:

1. how to facilitate technology transfer from research to practice;
2. how to foster research and infrastructure support for best practices in security and information fusion; and
3. strategies for funding research in this area.

Participants were invited from the areas of computer security, real-time systems, and information fusion, and included representatives from academia, industry, and government. Twenty people attended (see list at the end). The workshop was co-located with two workshops on Information Technology Research for Critical Infrastructure-hosted by UC Berkeley, Vanderbilt University, and the University of Virginia-in order to facilitate attendance at all events and to cross-pollinate ideas from different groups of people.

The aims of this workshop on research related to national security are directly relevant to CRA's mission. CRA seeks to strengthen research and advanced education in computing and allied fields. It does this by working to influence policy that impacts computing research, encouraging the development of human resources, contributing to the cohesiveness of the professional community, and collecting and disseminating information about the importance and the state of computing research. Each plays an important role in achieving the organizational objectives for the benefit of the country.

In the following sections, we discuss the recommendations that the workshop produced for each of the topic areas.

Research Areas

The workshop focused on three research areas-security, information fusion, and critical infrastructure-based on recommendations made by a National Research Council Committee on Science and Technology for Countering Terrorism. That committee called for research in information and network security, new information technology for emergency response, and new information technology for the detection, remediation, and attribution of attacks (information fusion).

Security includes authentication, availability, containment, detection and identification, privacy, recovery, and new security models. Information technology for emergency response includes a variety of problems- most notably problems for critical infrastructure. Information fusion includes research in data and text mining, data integration, language technologies, image and video processing, and evidence combination. Rather than consider again the question of which research areas are important for national security, the CRA Workshop took the results of the NRC committee's report as a starting point.

To provide common grounds for discussion, the workshop began with presentations by experts, who outlined the current state of the art and active research topics in each of these three areas.

Recommendations on How to Facilitate Interaction between Research and Practice in Security and Information Fusion

Top priority should be given to methods for facilitating interaction between research and practice. It is especially important that researchers have the ability to base their work on real problems with connections to real data. Because of concerns about national security and privacy, this can be particularly problematic. Nonetheless it is important that researchers and technologists have access to scenarios and data that are recognized as realistic and as representative of the challenges being faced by practitioners. If this is not the case, research results face the risk of being dismissed as irrelevant or immature.

In addition to providing access to the problems and the data, programs must be developed that facilitate an understanding of their counterparts by both researchers and practitioners. Researchers need a deeper understanding of the complex processes in which practitioners, such as government analysts, participate. They need to be able to observe practitioners and their processes in action. Practitioners need an understanding of the potential of new technology. Most are not comfortable with new technology, and novel methods for introducing technology must be developed so that people can become familiar with and test new systems-all while continuing to make progress on their real-world tasks.

Given these needs, the workshop recommended that the following actions be taken:

1. Create testbeds of open data. The workshop recommends establishing a center that will make it easier for both government and industry to provide data. In general, it is hard to generate synthetic data with enough scale; this is a research project worthy of its own funding. Issues include development of new techniques for automatic scrubbing, agreement between researchers and intelligence agencies on what constitutes good, normal operational and attack scenarios, and a long-term focus on establishing and maintaining the testbed. It was suggested that different research groups focus on different aspects of the testbed; one group might focus on generating operational data, while another focuses on generating attacks. Initial models for such a testbed are being explored under the NSF KDD program, a joint program with the National Security Agency. Funding for such models involving these and other agencies should be provided.

2. Establish structures that facilitate interactions. The workshop recommended a variety of structures that could address the problem of connecting research with practice. Grants focusing solely on the transfer of technology for a short-term period should be established. Funding programs that stress interactions between the intelligence community and industry and research groups are also needed. National laboratories that focus on issues of security and data mining would allow researchers and practitioners to come together for longer periods of time. Such laboratories could provide the ability to generate large-scale simulations in which experiments could be carried out. Organizations such as In-Q-Tel should seek to encourage technologies driven by needs and not by the market, with special effort placed on removing bureaucratic difficulties. Google provides a good model for moving from research to practice that could be used as the basis for new structures.

3. Adopt human factors methods for modeling and improving security processes. It was recognized that often the security processes that practitioners follow are cumbersome and error-prone. To facilitate understanding of the tasks and the human activities involved, research should consider and incorporate cognitive approaches such as scenario-driven exercises, workflow modeling, cognitive think-aloud protocols, and expert panels. Increasing the automation of many of these security processes, combined with rigorous analysis, would eliminate many opportunities for security breeches.

4. Reconsider research paradigms. Researchers and funders must look to long-term efforts that include the continual development and improvement of realistic testbeds and careful evaluation of technology based on those testbeds. The workshop recommended that multiple cycles of evaluation are needed. In the first cycle, researchers might work with end-users to see how they react to initial tool functionality and design. In later cycles, after responding to initial concerns, more rigorous evaluation could be undertaken. This is a process that may go through many cycles and takes time. Funding agencies and users must recognize that long time periods are needed for this process to work well.

5. Create measures of effectiveness. If practitioners are to understand which technologies are worth being deployed, they need measures of effectiveness that can help them distinguish and choose among options. Such measures should provide qualitative assessments of functionality and usefulness, as well as the more typical quantitative metrics.

Best Practices in Security, Real-Time Systems, and Information Fusion

The workshop addressed the question of best practices primarily through breakout groups that focused on each research area separately. It became clear, however, that there were commonalities across all areas. Unfortunately, it was agreed that there are not very many best practices within individual areas. It was difficult enough to define 'best practice,' let alone the appropriate problems for which best practices should be developed. Furthermore, best practices change so quickly that it would be difficult to create a static list.

Instead, the subgroups looked to mechanisms and processes that could be put in place to dynamically track best practices. We report on recommendations separately for each research area.

Information Fusion

While researchers are very often focused on tools and methods, we agreed that what needs to be disseminated and described to the more general community are the best tools for given tasks. We need a focus on the problem, not the tool. Thus, a summary of what different search engines do is not helpful; instead, practitioners need to know how it behaves in a specific context.

This subgroup recommended the development of a playground for tool evaluation. The playground would define scenarios and data against which tools could be tested. Such a playground might be set up on the Web, allowing researchers to post tools and practitioners to specify problems against which they could evaluate multiple tools. In order for this to work, researchers must agree on an annotation scheme for markup of data and common APIs for tightly coupled or distributed architectures. In addition to tasks, games should also be explored as a motivating mechanism for exploring the best fit of a tool.

In summary, this subgroup did not think it appropriate for any organization to develop a list of best practices; rather, it thought it would be better to define an environment for determining best practice, given a particular task. Best practice depends on context. This environment should be used to capture lessons learned. It should be developed as a glass-box scenario, logging behavior and allowing observation of end-users to see how well tools work, particularly when personal preferences play a role. It is possible that an organization such as the Linguistic Data Consortium at the University of Pennsylvania would be appropriate for setting up and maintaining such an environment, if provided with adequate funding.

Real-Time and Embedded Systems

The few best practices in existence include formal methods used for core algorithms, real-time analysis, and quality of service guarantees. In addition, there are common modeling and analysis tools in use, as well as integrated development environments. However, most of these tools are limited to idealized systems and situations. They also do not adequately address security and information fusion issues. Extensions to these tools and best practices are needed for all of these issues.

The most critical areas for which best practices are needed include methods to deal with the integration of constraints, dynamic real-time aspects of the system, dependable software development for real-time systems, computer security, and more principled development of large-scale distributed systems, which typically are still ad hoc.

This subgroup recommended the development of a set of Critical Infrastructure Protection (CIP) centers that focus on science and provide industry/research consortiums. Such centers could provide diversity on any given problem and will allow for integration of security with real-time issues. Different centers might focus on different problems-emergency response systems; wireless sensor networks for security of infrastructure systems for power, water, and transportation; or cyber security on the Internet-but cooperate with others.

Security

The security subgroup had the least agreement on what constitutes best practice, opting for the term "plausible practice" instead. Even security itself encompasses many possible areas, such as cryptography, network security, computer security, and security administration. The subgroup focused on security administration.

Recommendations include the need for more quantitative research on good security and evaluation. For improvement in security practice, the subgroup pointed out the need for creating novel forms of attacks on existing methods. Best practice is often limited due to the installed operating system and software, which are often decades behind the techniques put forth by the research community. This dichotomy between research and practice in security means that different recommendations must be developed for different situations. Given that all of our systems have vulnerabilities, it is unrealistic to expect that any system can ever be entirely secure. Instead, we need to move toward strategies that provide security components that are self-configurable and, in the case of attacks, self-healing.

Recommendations on Strategies for Funding Research in These Areas

There are a number of programs already in place at the different funding agencies to address issues of national security.[1]

The workshop recommends that a mix of funded programs targeting issues of national security be established and maintained. In particular, it is important that both short- and long-term efforts be supported; either type of effort alone is not sufficient. Four critical issues were identified as key to development of new technology for national security:

1. Improve mechanisms for funding technology transfer. We need better methods for funding efforts to deploy mature research into applications. Possibilities include 12- month funding augmentation at the end of existing grants or short-term grants focused entirely on technology transfer. Improvement of the Small Business Innovation Research (SBIR) model should also be considered. The conversion rate from Phase I to Phase II SBIR grants in the current model is fairly large, but many small companies are never weaned off of SBIR grants; when these grants end, the company also ends. A more gradual move between phases is needed. In addition to short-term efforts, a study of mechanisms that facilitate tech transfer is needed.

2. Establish support for longer-term research on national security. The problems will not be addressed by deployment of existing research alone. Many of the problems facing the intelligence community are hard ones and existing solutions are not available. Nonetheless, there are research efforts underway that are applicable to these problems that could be focused on this area. Funding programs that allow for the creation of centers and focused research over a long time period are needed. The need for open and realistic testbed data sources, comparable to the data used by the intelligence community, is one example of an area where new research is needed. These testbeds would in turn be used for other research. The NSF KDD and ARDA programs provide good models for this type of funding. Additional programs such as these in more areas are encouraged.

3. Create new programs that facilitate interactions between practice and research. Such programs could include a faculty center where faculty are given clearances, or a scholar-in-residence program where researchers spend a sabbatical or a shorter period of time at one of the intelligence agencies or national laboratories where researchers and practitioners could be brought together. Programs that embody cognitive methods for observing end-user needs and the use of demo and employed systems are particularly important. Flexibility and experimentation with new models for prototyping, testing, and redesigning systems are needed.

4. The research community must get involved. There is a need for more participation by the research community in funding programs. DARPA has a need for new program managers, and without them new programs will not be initiated. NSF also has a need for rotators who are willing to serve time at NSF to oversee funding programs. Research recommendations from the community are also influential in starting new programs at both DARPA and NSF. For example, in order to establish a cross-institutional workshop on a topic of relevance, NSF needs a White Paper from a university. Similarly, DARPA is open to suggestions from the community on new programs.

Workshop Participants

James Allan, University of Massachusetts, Amherst
Kelcy Allwein, Defense Intelligence Agency
Chris Buckley, Sabir Research
Jagdish Chandra, George Washington University
Yvo Desmedt, Florida State University
Helen Gill, National Science Foundation

Virgil Gligor, University of Maryland
Sally Howe, National Coordination Office
Andrew Hume, AT&T
Rob Kolstad, SAGE Executive Producer, USENIX
Jay Lala, DARPA
Carl Landwehr, National Science Foundation
Elizabeth D. Liddy, Syracuse University
Stephen R. Mahaney, National Science Foundation
Kathleen R. McKeown, Columbia University
Bert Miuccio, Center for Benchmark Services
Al Mok, University of Texas at Austin
Salim Roukos, IBM
Shankar Sastry, University of California, Berkeley
Jonathan Smith, University of Pennsylvania
John Stankovic, University of Virginia
Gary Strong, National Science Foundation

[1] The National Science Foundation (NSF) has created at least one program jointly with the National Security Agency under the Knowledge Discovery and Dissemination (KDD) program, and plans others. The mission of the Advanced Research and Development Activity (ARDA) is to work closely with the intelligence agencies and has several programs (e.g., AQUAINT, NIMD) where researchers and intelligence analysts are teamed to work together on problems and solutions. DARPA has no set-aside to address problems in information security, but initiatives can come through the program managers. That said, there are several ongoing security-related programs within DARPA that bring together research from different sites.