Realism in Teaching Cybersecurity Research: The Agile Research Process

Melissa Dark[1], Matt Bishop[2], Richard Linger[3], and Luanne Goldrich[4]

[1] Computer and Information Technology Department, Purdue University, West Lafayette, IN 47907 USA dark@purdue.edu
[2] Dept. of Computer Science, University of California, Davis Davis, CA 95616-8562 USA mabishop@ucdavis.edu
[3] Cyber and Information Security Research Group, Oak Ridge National Laboratory, Oak Ridge, TN 37831 USA lingerr@ornl.gov
[4] Johns Hopkins University Applied Physics Laboratory Laurel, MD 20723 USA Luanne.Goldrich@jhuapl.edu

Abstract. As global threats to information systems continue to increase, the value of effective cybersecurity research has never been greater. There is a pressing need to educate future researchers about the research process itself, which is increasingly unpredictable, multi-disciplinary, multi-organizational, and team-oriented. In addition, there is a growing demand for cybersecurity research that can produce fast, authoritative, and actionable results. In short, speed matters. Organizations conducting cyber defense can benefit from the knowledge and experience of the best minds in order to make effective decisions in difficult and fast moving situations. The Agile Research process is a new approach to provide such rapid, authoritative, applied research. It is designed to be fast, transparent, and iterative, with each iteration producing results that can be applied quickly. Purdue University is employing Agile Research as a teaching vehicle in an innovative, multi-university graduate program with government sponsor participation, as described in this paper. Because it simulates real-world operations and processes, this program is equipping students to become effective contributors to cybersecurity research.

1 A New Approach to Teaching Cybersecurity Research

Graduate programs in computer security emphasize research. Students who pursue a masters degree by writing a thesis or a doctoral degree by writing a dissertation are expected to do research that contributes to the body of knowledge. This dedicated study of a research problem requires students to be motivated and interested in the problem, and to understand the context in which the problem arises. The obvious way to do this is to study a particular facet of the problem in an applied circumstance. The student can then understand the constraints and available mechanisms, and attempt to develop a methodology or theory that will solve the problem in this specific instance. This leads to a good masters thesis. Doctoral students can then generalize the problem by relaxing the constraints

and examining the broader problem in other contexts. This leads to a traditional research problem, and hence to the dissertation.

Unfortunately, graduate students are often not exposed to research until they begin their thesis or dissertation work. The goal of the work described in this paper is to provide this exposure by integrating research into the curriculum. Specifically, it focuses on obtaining applied research problems from sponsors, and having the students apply research processes to the problems under the guidance of both faculty and sponsors. This assures the results will be of interest to, and usable by, the sponsors. It also leads to broader, more traditional research work that stems from the application.

This work aims to enhance students' ability to plan, organize, and carry out research, especially as a member of a team, under the combined mentorship of faculty and sponsors. Additionally, it provides a basis for educational institutions to give students research opportunities that bridge theory to practice by focusing on real-world problems with real-world applications. It does so through the mechanism of a class on computer security research, run in conjunction with industry and government sponsors.

Because of the limited duration of the class, the students cannot complete a full, long-term research project. But they can begin or continue one, and so the research must be organized in a way that produces deliverables of some sort for the sponsor within the time constraints of the class. Further, the students must carry out and document their work in such a way that a completely different team of students can pick up where the original students left off, and continue the work. A new research methodology called *Agile Research* [5] provides an ideal way to do this.

This paper describes this work. The next section discusses the basis for the class, its organization, and how the work proceeds. Then we present a review of Agile Research, and follow up with a description of how the class implements the phases of that research methodology. We conclude with plans for the future.

2 Applied Research Class Organization

One of the focal activities of the INSuRE (Information Security Research and Education) project [1] is an applied research class. INSuRE is a consortium of 10 universities (Purdue University is the lead institution, and the other participants are Carnegie Mellon University, Dakota State University, Iowa State University, Mississippi State University, Northeastern University, Stevens Institute of Technology, the University of California Davis, the University of Maryland Baltimore County, and the University of Texas Dallas), plus the U. S. Department of Defense, Sandia National Laboratory, Pacific Northwest National Laboratory, Oak Ridge National Laboratory, the Indiana Office of Technology, and Hewlett-Packard. INSuRE aims to develop a partnership among sponsors that perform cybersecurity research and need the results to perform their missions, and cybersecurity researchers who conduct the research and produce results, including students and faculty at Centers of Academic Excellence in Information

Assurance Research (CAE-R). INSuRE aims to become an agile, self-organizing, cooperative, multi-disciplinary, multi-institutional, and multi-level collaborative research project that can include both unclassified and classified research problems in cybersecurity. Currently, students from these 10 universities work on cybersecurity problems through coursework, directed independent study, and theses or dissertations.

The INSuRE applied research class provides an opportunity for students to work on problems provided by sponsors, as well as to be mentored by practitioners in the real world, rather than working solely on faculty-led research. More pressing and urgent problems are addressed, allowing the students to also benefit from the guidance of multiple and interdisciplinary research faculty from several institutions. The student-led research may in fact provide solutions for pressing national problems [7]. To facilitate scientific discovery, learning, and collaboration we use an open source software platform called HUBzero®. HUBzero includes a powerful content management system built to support scientific activities. Users on a hub can write blog entries and participate in discussion groups, but it is possible to do so much more. They can work together on projects, publish datasets and computational tools with Digital Object Identifiers (DOIs), and make these publications available for others to usenot as dusty downloads, but as live, interactive digital resources. Simulation/modeling tools published on a hub can be accessed with the click of a button. They run on cloud computing resources, campus clusters, and other national high-performance computing (HPC) facilities and serve up compelling visualizations.

Prior to the class, faculty solicit research proposals from external organizations in government and industry. These proposals are a paragraph or two in length, and describe a research problem in fairly general terms. For example, a proposal to examine biometric systems of authentication might be as follows:

> *Title*: Security of Biometric Authentication
> Biometric devices provide information about people that is often used to authenticate their identity. This information must be associated with other data that is used to match up the data from the device to the user. This raises two questions. First, how easily can the biometric device be fooled into reporting incorrect measurements? And second, can the user change the comparison data on the system? This project explores the second question by determining how to change the comparison data for a given biometric device.
> *Sponsor*: John Oldman
> *References*:
> - "Biometric Security", http://example.com/bio-security
> - Jacob Marley, "Attacking a Biometric System," *Journal of Christmas Past* **3**(1) pp. 1–20 (Jan. 1951).

This Spring, the list includes projects on forensics, using code variation as a defense, an analysis of the proposed TCPcrypt protocol, machine-assisted semantic understanding of code, profiling industrial control system nodes, and the

impact of known vulnerabilities upon layered solutions. The list is compiled and made available to faculty immediately and to students on the first day of the semester/quarter.

Obtaining sponsor interest has thus far been very successful; indeed, typically there are more proposed projects than there are students. Faculty members have solicited projects from people they know and, in many cases, have worked with. Most projects have come from government groups, but industrial firms and organizations have also proposed several. Interestingly, the latter typically take longer to prepare and get approval for projects than do the government organizations. For example, at least two companies were hoping to propose projects for the Spring term, but were unable to obtain the necessary approvals in time. Whether students can propose their own projects is up to the faculty member teaching the class. Some faculty allow this if the projects are substantial enough and deal with a current topic, on the basis that the students are best motivated when they are working on a project that they feel strongly enough about to propose. Other faculty members prefer that students select from the sponsor projects. As the proposed projects are typically broad, the students and sponsors have had no trouble narrowing down the projects to be of interest (sometimes enthusiastically so) to the students.

Complicating project selection is that different universities have different rules about working with sensitive projects. For example, the University of California at Davis does not allow any classified work to be done on campus, because that would restrict the ability of the researchers to publish (among other reasons). But other universities do. Proprietary work for industry has similar but different constraints. Thus, all sponsors must agree that, should the results and the work merit publication, the research from any project they propose can be published. As of now, this has not been a barrier to obtaining interesting projects.

The students prepare bids on at least two projects. First, the sponsors make a brief presentation to clarify their research needs and goals. Then the students engage in exercises to identify the knowledge, skills, and competencies required to work on the projects they are interested in. Each bid has four key components: a personal statement of interest, a description of the research problem (the most substantive section), the expected outcomes, and a description of student's skills, knowledge, and abilities relevant to the problem. Based on the students' bids, the faculty and research sponsors move quickly to form research teams.

Critical to the success of the project, of course, is that the team members work well together. In some cases, faculty and sponsors select the students that make up each team, which requires judging how amenable the members are to one another. The rationale for pairing students in teams can be based on student interest, expertise, and/or work style. In other cases, the students organize their own teams; the faculty and sponsor must accept the membership. Having students organize their own teams provides them with an opportunity to consider the factors that will constitute a research team, which is a valuable lesson. This ensures that the teams are balanced. Sometimes team membership

changes after the initial formation. For example, at the University of California at Davis, a team of three members was reduced to two because one of the students became more interested in a different project, and so moved to the other team. However, all the students knew one another to some degree, and there were no problems with the change. The project has also started forming cross-institutional teams for the first time in Spring 2015. There is a three-person team with students from Purdue University and Mississippi State University, and another two-person team with a student from Dakota State University and one from Purdue University.

These teams next prepare a proposal, the contents of which are similar to that which would go into a National Science Foundation proposal (but with much less detail). The key components of the proposal are the review and analysis of previous work, and the statement of the specific aims of the project. The proposal also contains a schedule of milestones that the students believe they can meet, a plan describing how the students will approach solving the problem, and a bibliography. It also requires a realistic schedule and budget, a list of deliverables, and a discussion of any foreseeable difficulties and anticipated plans to overcome them. When writing the proposals, the students interact iteratively with sponsors and faculty to define the scope of the problem and near-term action steps to be taken. This step is critical in helping students assume the research problem as theirs as opposed to a work for hire, where the sponsors have "dictated" the scope of work and the students are simply following directions.

Once the proposal is approved, the students begin their research. As a first step, the students conduct a thorough literature review. This augments the quick literature reviews done earlier. Those reviews are simply aimed to show that the project has not been done earlier, and that it is substantial enough to advance the state of cybersecurity in some way. This literature review is structured around an argument or arguments. Typically, these arguments point out critical gaps in the existing literature, or how the work in that literature might be extended. If the goal of the research is to validate or correct a published result, the argument would explain the context of the work to be validated or corrected, why it is important, and what would happen if the prior results were incorrect or not corrected. The literature review is of sufficient importance that it is treated as an assignment and is weighted as much as the proposal is weighted.

Following that, the students begin acting on the plan laid out in the proposal. However, the teams are not left on their own to simply execute a 10-week project plan. Instead, teams meet every week with faculty and sponsors to report progress, the challenges encountered, how they are dealing with those challenges, and the next weeks goals. The goals sometimes change based on the challenges encountered. The rapid, successive iterations permit sponsors to modify incremental research goals and apply results based on intermediate findings as the work progresses (the principle of incremental management within a semester or quarter), and allows students to experience first hand the truly iterative and fast-paced nature of cybersecurity research.

The class requires students to prepare a midterm progress report that is delivered as a formal presentation to all classes across the universities via teleconference, and a final project presentation that also incudes a written report and poster. At the end of the semester, all students present the results of their research. For those on a semester system, this is a final presentation and report. For those on a quarter system, this occurs in the middle of the second quarter, and so is a penultimate presentation and report. Finally, those on the quarter system do a final project report and presentation at the end of the second quarter. The sponsor and the faculty member then evaluate the project.

One critical aspect of the final assignment is to document the progress made in a manner that allows the sponsor to iterate the next increment, and allows a team (at the same university or a different university) to pick the project up the next semester/quarter and continue the research, also in an agile, iterative manner. The specific manifestations of this differ based on the nature of the project. In some instances it includes a theoretical model that is sufficiently explained to allow a new team of student researchers to simulate and test the model. Another instance might be curating a dataset in a manner such that it is available for reuse and preservation. This type of documentation is essential to enabling the sponsor and faculty to incrementally manage research projects across semesters or quarters, and across institutions.

The sequence below summarizes the steps that the students follow [2]:

1. Project bid
2. Project proposal
3. Literature review
4. Progress report and presentation
5. Final report and presentation for schools on semester system; penultimate report and presentation for schools on quarter system
6. Final report and presentation for schools on quarter system

Given that the class uses a non-traditional model of research, specifically one with a much tighter time-line than traditional research, a new model is needed. Fortunately, such a model has been developed: the Agile Research process.

3 Agile Research

Traditional, long-term research often involves extensive requirements definitions, comprehensive proposals, competitive awards, distributed organizational structures, complex funding protocols, and long-term performance that can extend for years. When the scope and scale of research requirements are large, these traditional processes and their management procedures are essential to maintaining control across collaborating organizations and reducing risks of overruns and non-performance. As such, they serve a vital role in conducting large-scale, long-term research projects to achieve national goals [3, 7].

But events occur in cybersecurity areas that require fast and decisive responses in order to protect national well-being and even survival [4]. These responses would benefit from rapid and authoritative analysis by the best minds

and organizations. The traditional research infrastructure was never intended for this level of fast engagement and immediate application, and is not well suited for these situations.

The Agile Research process [5, 6] was developed to address the need for fast and effective researcher participation in situations where speed is an overarching requirement. When attempts have been made to apply traditional methods in these situations, the research results are often too late to be of use in the current cybersecurity event, and wind up sitting on a shelf, unused and forgotten.

Agile Research is organized around sponsors, who pose research questions to be answered, and researchers, who conduct the research and produce results. Sponsors and researchers may be in the same or different organizations, and may be organized in any number of ways provided the following principles are satisfied [5, 6].

- *Predefined Infrastructure Principle*: Resources and logistics must be predefined and allocated before research needs emerge, to permit immediate deployment for fast engagement when needed. Agreements between sponsors and researchers regarding organizational roles, research capabilities, and contracting, funding, and intellectual property must be in place and ready to be instantiated in unforeseen circumstances with no delays. This "load-and-go" approach permits fast reaction by pre-positioned resources to unpredictable research needs unburdened by logistical constraints.
- *Incremental Research Principle*: Agile Research is structured into iterative, short-term, accumulating increments that each produces actionable results. Increments must first focus on understanding the problem, progress to solution strategies, and then to incremental solutions. Understanding how to organize research into a series of accumulating, referentially transparent increments requires careful planning. Early increments must provide a framework for inserting and composing later increments such that results accumulate with little or no revision of prior work.
- *Incremental Management Principle*: The incremental research process provides built-in, short-term checkpoints for sponsors to understand researcher progress, and to direct subsequent work based on incremental findings. Agile Research projects can be quickly refocused based on changes in both fast-paced problem environments and on intermediate shortfalls and windfalls in the research. Visibility, transparency, and clear communication between researchers and sponsors are essential for informed management decision-making.
- *Transferability Principle*: Agile Research projects may be carried out by one group of researchers, but ready transfer of results from one group to another must be possible if necessary. As research increments are completed and changes in direction are made, mechanisms for quickly repositioning the research and resources to a new team must be in place. This includes knowing where the research expertise exists for the next increment, as well as providing supporting documentation that permits a new team to pick up the work seamlessly and rapidly.

Agile Research projects proceed through up to four stages, each culminating in researchers delivering results, either through briefings, white papers, tools, or a combination of these. At the completion of each stage, the sponsor decides whether and how to proceed. This process is summarized in Figure 1.

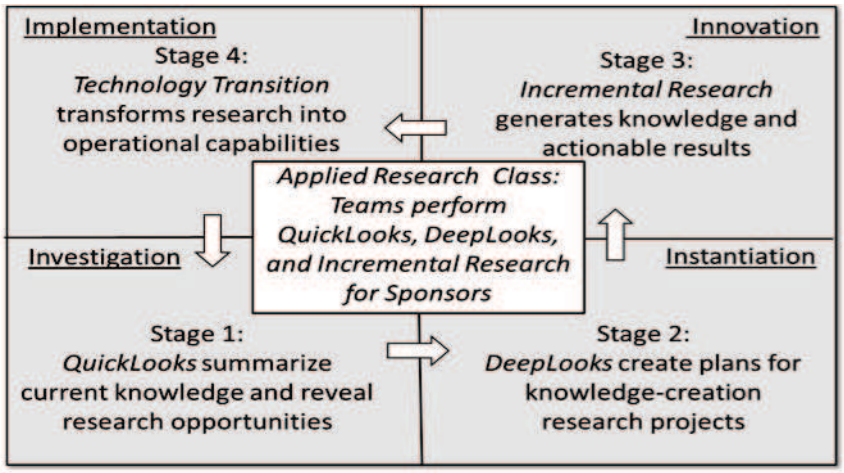

Fig. 1. The Agile Research process

- *QuickLook Stage*: This first stage generally takes days or weeks. It answers the question of what is known now about the problem. The research team clarifies research needs with the sponsor, explores the existing knowledge base, identifies subject-matter experts (SMEs), and provides recommendations to form a foundation for the research effort. This stage is deliberately made flexible to accommodate urgent or even emergency needs.
- *DeepLook Stage*: This second stage generally takes weeks. Based on results from the QuickLook stage, it answers the question of what the research can be expected to accomplish and how should it be done. It defines the research goals and plans in terms of iterative, accumulating increments that produce useful results for sponsors.
- *Incremental Research Stage*: This stage consists of multiple incremental steps, generally performed in weeks or months per increment. Each iteration adds to an evolving solution to the problem. This step-wise approach permits sponsors to modify incremental research goals and apply results based on the intermediate findings.
- *Technology Transition Stage*: Finally, if a project requires technology transfer, this stage, generally performed in months, provides specifications, prototypes, and support to guide technology implementation and operational use.

Agile Research is flexible. A project might require only a QuickLook to determine the state of knowledge for a particular problem. Or, a project could continue to a DeepLook to understand what the research could accomplish were it continued to the next stage, and how the research in that stage should be structured. The sponsor could then initiate the incremental research.

4 Putting the Class and Agile Research Together

The Agile Research model is well suited for the INSuRE class. Its structure corresponds closely to the first three phases of that model: the QuickLook Stage, the DeepLook Stage, and the Incremental Research Stage. Figure 2 depicts integration of the Agile Research process with the applied research class and the sponsoring organizations.

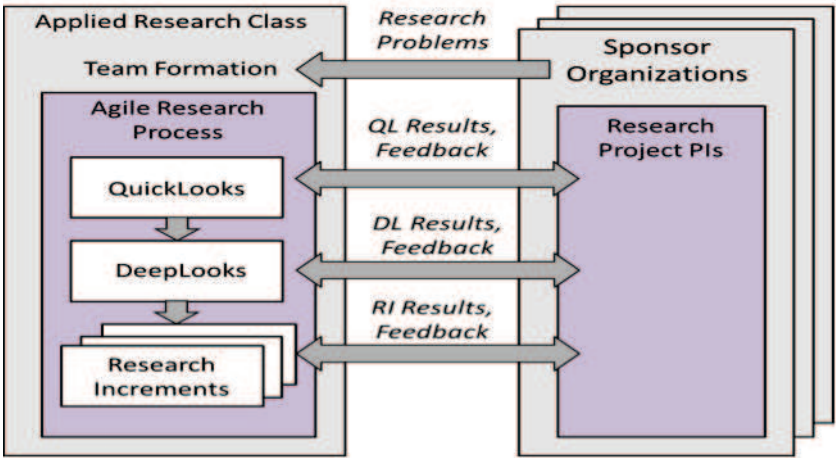

Fig. 2. Integration of Agile Research and applied research class

In the first stage of the class, the students do background work to prepare their bids and begin to scope the project. The sponsor starts the process by proposing a rather general research problem. Then each student makes a preliminary exploration of related work and decides on a view of the research project that he or she would like to explore. These recommendations will form a foundation for the research should the bid be accepted. In addition, the student identifies the competencies necessary by explaining why he or she is qualified to carry out the research.

This matches the QuickLook Stage of the Agile Research process almost exactly. The single difference is in the identification of the subject-matter experts. Rather than identifying others who are already these experts, the students explain why they should be considered, or will become, the subject-matter experts.

In a non-class setting, the subject-matter experts may well not be the people performing the QuickLook stage. The timing also matches. The bidding process for the class takes between 1 and 3 weeks; the controlling factor is the availability of the sponsors. In a non-class setting, the QuickLook would take about the same amount of time, again the availability of the sponsors being a critical factor.

The DeepLook Stage corresponds to the preparation of the proposal. Based on the bids, the students, faculty, and sponsors form the research teams. The teams then prepare proposals, as described above. The proposals present the goals of the research project, just as the DeepLook requires an answer to the question of what the research can be expected to accomplish. It contains a plan, saying how the research is to be done. Thus, it matches the DeepLook phase exactly.

The research itself instantiates the Incremental Research Stage of the Agile Research process. As noted above, the students meet with the sponsor weekly, with specific goals being set each week, and based on the results of each weeks progress, the sponsor can modify the research goals. Further, the sponsor can apply intermediate results from the teams work. This matches the goals and design of the Incremental Research Stage.

The design and implementation of the class reflects the principles of Agile Research. Of particular note is the transferability principle, which says that the results of one group must be transferable to another group. This is exactly how each team wraps up its results at the end of the class, because another team, possibly from another university or even split across multiple universities, may choose to pick up where the research was left off. Similarly, the incremental management principle requires that the research progress be incremental, with checkpoints for the sponsor and team to confer and determine how best to proceed; the sponsor will also receive actionable results at each increment. Again, this is reflected in the weekly meetings between the team and the sponsor.

5 Conclusion

The INSuRE program was begin over a year ago. Initially four universities were involved; the success of that initial year encouraged six more universities to join, and more organizations to propose problems. The Agile Research model was developed for a different purpose. However, it very closely mimics the desired approach used in the class, so applying the model provides a framework to support the effectiveness of the research process used in the class.

The use of Agile Research in this context raises some interesting questions. The work performed here is public (not classified nor proprietary), because some of the universities require that any research conducted must be publishable. How would this model work in a non-public arena? Would the proprietary or governmental constraints interfere with the educational benefits of applying the Agile Research model?

Another question is measuring how effective this approach is in training the students in research methodology. Can the techniques they learn here help them conduct the more traditional, long-term research needed for a doctoral dissertation? The intuitive answer is yes, because the structure of planning the research to produce publishable intermediate results will provide the students with a strong publication record when they complete their dissertation research, and the intermediate results may cause them to refocus the research if those results indicate the expected results will not hold or cannot be done. But the nature of the research—applied vs. pure—may pose a clash in the two approaches. Agile Research begins as a very applied research methodology, but traditional academic research is intended to extend the body of knowledge in ways that may not be immediately applicable. As an example, Riemannian geometry was developed in the 19th century as a demonstration that Euclid-s fifth postulate was in fact an axiom and not a provable proposition. It had no realistic applications until the 20th century, when the geometry of the universe was found to be Riemannian and not Euclidean. Were the goal of the research to develop a useful geometry, Riemannian geometry might never have been developed.

There is of course a place for both applied and pure research—indeed, pure research often provides the tools upon which applied research builds, and applied research often motivates the questions that guide pure research. Agile Research, with its emphasis on actionable results, is more applied, but leads to the fundamental questions that students can examine in their dissertations. Thus, it fills an important niche, and when used in an educational setting such as the INSuRE class described here, provides a firm foundation for students to begin a successful cybersecurity education and career.

Acknowledgements: Melissa Dark and Matt Bishop were supported by the National Science Foundation Grant Number DUE-1344369 to Purdue University, and by a subcontract from Purdue University to the University of California funded by that grant. Matt Bishop was also supported by the National Science Foundation Grant Number OCI-1246061 to the University of California at Davis. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation, Purdue University, or the University of California.

Richard Linger worked on this manuscript as an employee of UT-Battelle, LLC, under Contract No. DE-AC05-00OR22725 with the U.S. Department of Energy. The United States Government retains and the publisher, by accepting the article for publication, acknowledges that the United States Government retains a non-exclusive, paid-up, irrevocable, world-wide license to publish or reproduce the published form of this manuscript, or allow others to do so, for United States Government purposes. This submission was written by the author(s) acting in their own independent capacity and not on behalf of UT-Battelle, LLC, or its affiliates or successors.

References

1. INSuRE eager (2013), http://www.nsf.gov/awardsearch/showAward?AWD_ID=1344369

2. Ecs 289m spring quarter 2015: Introduction to research in computer and information security (2015), http://nob.cs.ucdavis.edu/classes/ecs289m-2015-01/index.html
3. Branscomb, L.M., Auerswald, P.E.: Between invention and innovation an analysis of funding for early-stage technology development. Technical Report NIST GCR 02–841, National Institute for Standards and Technology, Gaithersburg, MD, USA (Nov 2002), http://www.atp.nist.gov/eao/gcr02-841/contents.htm
4. Fonash, P., Schneck, P.: Cybersecurity: From months to milliseconds. IEEE Computer 48(1), 42–50 (Jan 2015)
5. Linger, R., Goldrich, L.: Agile research for cybersecurity. Tech. rep., Institute for Information Infrastructure Protection, Dartmouth College, Hanover, NH, USA (Jun 2014), http://www.thei3p.org/docs/research/agile08-2014.pdf
6. Linger, R., Goldrich, L., Bishop, M., Dark, M.: Agile research for cybersecurity: Creating authoritative, actionable knowledge when speed matters. In: submitted for publication (2015)
7. Maugham, D.: The need for a national cybersecurity research and development agenda. Communications of the ACM 53(2), 29–31 (Feb 2010)