Computer Forensics at the 2006 Alaska Summer Research Academy

Christopher Hecker, Kara L. Nance, and Brian Hay, *ASSERT Center, University of Alaska Fairbanks*

Abstract* – A computer forensics course was offered during the 2006 Alaska Summer Research Academy (ASRA) at the University of Alaska Fairbanks. The two-week course provided a small number of high school students with the opportunity to gain experience in and an understanding of the field of digital forensics. Topics covered in the course included ethical issues related to digital forensics, digital footprints, forensics for digital media, and network-based forensics. It also included presentations by UAF instructors, guest lecturers, and the students themselves. While the 2006 course was very successful, the lessons learned during that session will be used to improve further offerings in this area, both within the ASRA program, and in the wider educational and outreach mission of the university.

Index terms – Computer Forensics, Digital Forensics, Alaska Summer Research Academy, ASRA.

I. The Alaska Summer Research Academy

The Alaska Summer Research Academy (ASRA) is an annual academic experience offered by the College of Natural Science and Mathematics at the University of Alaska Fairbanks (UAF) aimed at middle and high school students. In recent years ASRA has hosted approximately 140 students each summer, and while the majority come from schools within Alaska, there are also many students who travel from around the United States to participate [1]. On their application, participants select a topic, and then spend their thirteen days at ASRA involved in an in-depth study of their chosen subject, led by ASRA faculty and staff. In 2006, the sixteen available fields of study included topics such as marine biology, photography, robotics, wildlife ecology, and for the first time, computer forensics. While eight hours each day are devoted to educational experiences, ASRA is a residential program, and students from multiple disciplines have the opportunity to meet and exchange their educational, personal and cultural experiences. Students typically report that ASRA is an enjoyable experience, and the program commonly sees students participating in multiple years, during which they study a variety of subjects.

II. ASRA Computer Forensics Module

A Computer Forensics module was offered as part of ASRA in 2006, and is being offered again in 2007. Computer forensics was the first ASRA module in 2006 to reach capacity, (eight students based on the physical limitations of the computer lab being used). In addition, the computer forensics also had a waitlist of 16 students who could not be accommodated in 2006. This initial offering was very successful, and the lessons learned by the ASRA faculty and staff during 2006 will be used to enhance the experience for students in the 2007 session.

III. Course Naming

The term "computer forensics" has quickly become outdated, in that investigations commonly involve far more than those devices traditionally thought of as computers. For example, digital media players, cell and IP phones, PDAs, portable mass storage devices, gaming consoles, and network devices are all commonly vital targets of "computer forensics" investigations. As such, the term "digital forensics" is currently accepted and preferred in academia and by practitioners, due to its inclusion of all manner of digital devices. However, this term has not become part of the public lexicon, and as such it was felt that naming the module "Computer Forensics" rather than "Digital Forensics" would be more meaningful to the target audience.

*ASSERT Center, Department of Computer Science
University of Alaska Fairbanks
210 Chapman
Fairbanks, AK 99775-6670
http://assert.uaf.edu
assert@uaf.edu*

A. *Common Perceptions Of Digital Forensics*

The concept of digital forensics as a component of investigations has entered the public consciousness largely as a result of its portrayal in popular television shows, such as CSI, Law & Order, and 24. However, these shows typically present the topic in a highly unrealistic light, both in terms of the timeframe required for an investigation, and in the capabilities of the investigators and their tools.

B. *Student Background and Experience*

The 2006 group consisted of 8 students, ranging in age from grade 8 to grade 10. More important than age difference, however, was the range of computing experience in the class. At one end of the spectrum, some students had programming or scripting experience, and understood basic operating system and networking concepts. However, other students essentially had essentially just used computers for email, messaging, browsing, and gaming. In addition, the large majority of students had limited, if any, experience with operating systems other than Windows. This wide variety in abilities made the development of appropriate lessons extremely challenging, and in such a short session there is very little opportunity to provide remedial tutorials to ensure that all students have a base proficiency level.

In 2006 the computer forensics module had 24 applications for 8 places, and it is likely that a group with more similar skill levels could have been selected had the application process included a more in-depth evaluation of the applicants computing background. The option of offering the subject as two independent modules, one of which is appropriate for students with more basic computing skills and the other offering advanced topics for those with more experience, has also been considered for the 2008 session, should the interest in the topic continue in 2007. Currently for the 2007 session, we are examining the skill set of the students more closely in the application process, and explaining the amount of time spent inside the lab to give the students a better understanding of the module and associated expectations. Though there is some concern that limiting the module to all advanced students might not bring in a diverse enough which might actually be a detriment to the module.

IV. Topics Studied In The 2006 ASRA Computer Forensics Session

The 2006 session involved the study of several topics related to digital forensics, and great care was taken to ensure that the topics were presented in a manner appropriate to the maturity level of the students involved in the program.

A. *Digital Footprints*

The first activity in the session involved the students exploring the UAF campus in an attempt to determine where and how they left a digital footprint while performing everyday activities. Examples that the students found included the use of student ID cards to purchase food or checkout library books, credit and debit cards at campus stores, computer use in the campus labs, and video surveillance. The students then regrouped for a discussion of the exercise, and a further investigation of digital footprints beyond the campus environments, such as cell phone records, electronic highway toll systems, and Internet browsing records. This activity provided students with understanding of the extent to which digital evidence is left as a result of many common activities, and also served as a clear example to the group of why digital forensics was a more appropriate term than computer forensics.

B. *Ethics*

Although the digital forensics techniques and tools presented in the ASRA course are meant to be applied in the context of authorized investigations, there is always the potential that they can be used inappropriately by students. As such, it is important that such courses include an ethics component, and this was indeed a major part of the ASRA course, beginning even before the session started. The application process for computer forensics was more rigorous than that for other ASRA modules to increase the likelihood of selecting only those students who were sufficiently mature to be exposed to digital forensics techniques. As part of their application packer, each student was required to write and submit a one-page essay in which they described some of the ethical issues involved in the study of digital forensics. Furthermore, each student underwent a brief telephone interview, and the letters of recommendation from teachers were required to specifically address whether the applicant possessed the appropriate level of maturity and responsibility necessary for study of such a field.

Ethics was not only a component of the ASRA application, however. A full day at the start of the course, immediately following the digital footprints exercise, was devoted to the ethical aspects inherent in the digital forensics field, and also in the use of digital systems in a more general sense.

The requirement for each student to use their knowledge only in appropriate ways was reinforced each day, and it was made clear to students from the beginning of the course that students who violated any rules regarding the use of tools or techniques would be immediately removed from the course. The first day of camp the students were required to sign the Advanced System Security Education, Research, and Teaching (ASSERT) Lab Usage Policy, since this was the environment in which the module was to take place.

C. Forensics for Digital Media

Much of the first half of the course focused on forensics techniques and tools related to digital media, including hard disk drives, USB mass storage devices (thumb drives or memory sticks) and flash cards. The process started with a basic discussion of filesystems, which gave the students an understanding of how data can be organized, how files are typically "deleted", and why there can be much more data available for recovery than may be reported by tools such as Windows Explorer.

Once the students had an understanding of the relevant theory, they then had the opportunity to experiment with various methods for forensics investigation of digital media. The initial and most basic tools included *dd*, *strings*, and *md5sum*, and used the Linux operating system (although these tools are also available on a variety of other systems, including various Windows operating systems) [3, 4, 5]. During the initial exercise students were given a pre-configured floppy disk on which there were several valid files, and some which had been deleted. They initially viewed the disk using the standard Linux tools, such as *ls*, to see the files that the operating system reported as being on the disk. The students then made an image of the disk using *dd*, and verified that the image was a valid copy of the original disk by computing the MD5 hash values for the disk itself and the image they created. Finally the *strings* command was used to search the disk image for ASCII data, resulting in the recovery of text from files which had not been visible through the standard operating system tools.

The next exercise involved the search for photographic images, specifically JPEGs, on a flash drive taken from a digital camera. Each pair of students was given a digital camera, and sent out to take a series of photographs. They then deleted some of these photographs from the camera, and then took some additional photographs before returning to the lab. The groups then exchanged cameras, and attempted to recover as many images as possible from the flash drive using the Linux dd and strings commands. In the past a similar exercise has been used in programming classes at UAF, in which students are required to write a tool which searches for the start and end of JPEG files on a forensic image of a flash drive, and then attempts to recover the original JPEG images based on those markers. However, due to the lack of programming experience of several students it was felt that the programming component of this exercise would not be productive in the ASRA course. The instructor did, however, discuss how such a program could be constructed, based on the definition of the JPEG file format.

The last exercise in the digital media section of the course focused on the use of digital forensics tools such as Encase and the tools found in the Backtrack Linux distribution [2]. The UAF ASSERT Lab has several Encase installations, in addition to four stations equipped with FastBloc hardware write blockers, and the students were given pre-prepared 1GB hard disk drives [6, 7]. The goal of the exercise was to recover information planted on the disk in a variety of ways, forming a digital scavenger hunt that eventually allowed the students to solve a puzzle.

Although the process required to conduct legally sound forensics examinations was not a major focus of this course, some of the relevant concepts were discussed, and this was the case during this component. The concept of performing an investigation without making modifications to the evidence was discussed in the context of an examination of a hard drive, as were the reasons for using a hardware write-blocking device as opposed to hoping that the operating system will not modify a disk which is attached in a more traditional fashion.

D. Network Based Forensics

The varied backgrounds of the students made a basic network primer session necessary, as many of the students had very little computer networking knowledge whatsoever. This focused on giving the students a basic understanding of how Ethernet, TCP, UDP, and IP work, allowing them to participate in the network forensics sessions. Students were also introduced to the common hardware components found in computer networks, such as switches, hubs, and routers. Clearly much more time could have been spent on these issues alone, and some of the students already had a deeper understanding of networking concepts, but it was felt that this basic networking primer would at least allow all of the students to understand how digital forensics could be applied to a networked environment, which was the true purpose of the course. The primer section culminated in the students building small scale networks in the ASSERT Lab, using hardware components and pre-built virtual machines.

There was some concern that the network based tools and techniques were the most likely to be abused by the students, but network traffic is increasingly important part of forensic investigations, and it was felt that this was important topic that could be incorporated into the program if done carefully. The network-based sessions were conducted on the ASSERT Lab's isolated networks, allowing even those students who had little networking experience to experiment without fear that their actions could impact any production environment. The ethical issues associated with performing these activities in any other environment were stressed again throughout this session, as the misuse of networking tools is the easiest way for students to find themselves facing criminal or civil charges and possible expulsion in an academic environment. The placement of the network based exercises later in the sessions also gave the instructors time during the first week of the course to determine if any of the students were not sufficiently mature to accept the responsibility associated with the network forensics material.

After the completion of the basic networking primer, the session continued with a demonstration of how various types of common traffic, such as HTTP, HTTPS, POP, SMTP, Telnet, and SSH, appears on the network, and what information can be gathered in each case. This was an enlightening experience for many of the students who were not aware of the extent to which much of their Internet traffic, including email messages, is essentially available and accessible to anyone who cares to look at it.

The network based session also reviewed some commonly accessible tools for network forensics, including *tcpdump* and *Wireshark*/*Ethereal* [8, 9]. Great care was used in exposing students to these tools, and the session began with a repeated warning that the use of such tools outside the lab would likely result in very serious consequences. In addition, capturing packets on modern switched networks typically requires some additional effort to force the switch to send traffic destined for other hosts to the monitoring systems, and students were given no information about how that could be accomplished, or even that it was necessary. In addition, the packet analysis exercises conducted by the students used pre-captured data, rather than allowing the students to capture packets from the network themselves.

An example of a network based image consisted of a students being given a packet capture which, in part, consisted of web traffic. Students were asked to identify the sites visited during the capture period, and also to reconstruct some of the images that were downloaded during the session.

E. Guest Speakers

While Alaska does not have an extensive IT industry, it was possible to incorporate guest lectures by three law enforcement agents whose work, to varying degrees, involves the use of digital forensics. Prior to each speaker's presentation the students discussed the areas that they were interested in learning about, which gave the presenters a range of topics from which to draw, and ensured that the presentations were both relevant and interesting to the students. The first presenter was an officer with the Fairbanks Police Department who was new to the digital forensics field, having only recently completed training in that area, and his talk gave the students an understanding of how digital forensics is being taught in the police force.

The second speaker was an officer with the UAF Police Department who has extensive digital forensics experience, and often provides digital forensics assistance to law enforcement agencies and departments around Alaska. His talk focused on the types of investigations and evidence that he commonly sees, and also how some of the tools, such as Encase, are applied in his work.

The final guest speaker was an FBI agent stationed in Fairbanks. Although the FBI does not have a dedicated digital forensics examiner in Fairbanks, due to the small size of the community, the agent was the liaison with the FBI in Anchorage for digital forensics issues, and was also able to discuss the practical applications of the techniques and concepts during the course.

F. Student Presentations

The students were required to perform independent research and present their findings to the group on two occasions during the session. The first presentation involved the computer worms and viruses, and how digital forensics techniques can be applied to detect the presence of a new worm or virus then determine what the impact of the malware on a system. The second presentation focused on encryption and steganography, and how they impact digital forensics investigations, including how their use can be detected, and what information can be gathered despite their use.

V. Lessons Learned

The 2006 Computer Forensics class was by most accounts very successful, but improvements can certainly be made for future offerings. One major issue was the students' expectations and initial understanding of the subject, which was driven in large part by Hollywood's glamorization of the field. Of these misconceptions, perhaps the greatest was the time frame required for a typical investigation: on television systems are analyzed in minutes, whereas the reality, as was highlighted by both the practical exercises and the guest speakers, is that investigations can take days, weeks, or months, and in many cases there is just not sufficient time to perform an exhaustive examination. Even the exercises which featured unrealistically small datasets (such as hard disks with a 1GB capacity) required a significant amount of time to analyze. To address this issue an attempt was made to incorporate puzzles or scavenger hunts into the data, thereby holding the interests of the students for longer periods.

Another issue was the need to be inside working in the computer lab for much to exercises, which is particularly galling to Alaskans who already spend much of the year indoors, and value their short summer when they can spend time outside. This was addressed to some extent addressed by sending the students out of the lab when possible, such as when taking pictures with the digital cameras or finding examples of the digital footprint they leave.

The students did enjoy the change of pace that was offered by the PowerPoint presentations. These activities allowed them to research a facet of a topic which interested them, at a level that was relevant to their abilities and experience, and then present their results to their classmates. In addition, it provided an opportunity to incorporate written and oral communications skills, which have relevance to many aspects of their lives, into the ASRA session.

The guest speakers were also very successful in that they both gave the students a look into the real world of digital forensics, and held the students' attention by addressing the issues that they had expressed in an interest in. To supplement the guest lecturers, we included some videos presentations from conferences, but these were not received particularly well by the students.

A major issue that had a real impact on the course was the wide variety in experience of the students, particularly in the networking components of the course. The result of this was that some students finished activities far more quickly than others, and the classroom degenerated into individual session to a certain extent, which was very challenging for the instructors. In some cases, however, this issue was alleviated by pairing up students with varied experience levels, but a more even skill level would make the class as a whole significantly easier to plan and manage. In addition, such a scenario would ensure that sessions such as the networking primer were relevant to the entire group, rather than some students struggling to absorb the new information, while others sat through material that they already understood. An improved application process, or even a pretest of some form, may be necessary to address this issue.

VI. FUTURE PLANS

The Computer Forensics modules will be offered again, and there are several other topics that could be incorporated, particularly with a student body which had a more uniform level of experience. For example, neither the legal aspects of digital forensics nor the process by which a first responder or digital examiner should treat an examination site were addressed in any real depth, but these are important topics for a comprehensive understanding of the topic.

In 2006, no attempt was made to approach the topic from a programming perspective, which could be interesting for some groups of students. In addition, the issue of open research problems in the area of digital forensics, such as the problems associated with cataloging and analyzing the typically huge volumes of digital evidence, was not covered in this course, although this would make an excellent subject for an advanced course with an experienced group of students.

VII. CONCLUSIONS

The 2006 ASRA digital forensics class was a very successful first effort, and demonstrated that this topic could be presented in a manner appropriate to pre-college students. However, improvements can and will be made before the course is offered again in summer 2007, particularly in the application process in an effort to get a group of students whose background computer skills and experience is more homogenous than that of the 2006 group.

VIII. REFERENCES

[1] *Alaska Summer Research Academy*, retrieved March 11th 2007 from http://www.uaf.edu/asra/

[2] *Remote-Exploit.org*, retrieved March 11th 2007 from http://www.remote-exploit.org/backtrack.html

[3] *Linux Man Page – dd*, retrieved March 11th 2007 from http://linuxmanpages.com/man1/dd.1.php

[4] *Linux Man Page – strings*, retrieved March 11th 2007 from http://linuxmanpages.com/man1/strings.1.php

[5] *Linux Man Page – md5sum*, retrieved March 11th 2007 from http://linuxmanpages.com/man1/md5sum.1.php

[6] *Advanced System Security Education, Research, and Training Center*, retrieved March 11th 2007 from http://assert.uaf.edu/

[7] *Encase Forensic – md5sum*, retrieved March 11th 2007 from http://www.guidancesoftware.com/products/ef_index.aspx

[8] *Linux Man Page – tcpdump*, retrieved March 11th 2007 from http://linuxmanpages.com/man1/tcpdump.1.php

[9] *Wireshark: The World's Most Popular Network Protocol Analyzer*, retrieved March 11th 2007 from http://www.wireshark.org/