Mobile Phone Forensics Tool Testing: A Database Driven Approach

Ibrahim M. Baggili
Richard Mislan
Marcus Rogers
Purdue University

Abstract

The Daubert process used in the admissibility of evidence contains major guidelines applied in assessing forensic procedures, two of which are testing and error rates. The Digital Forensic Science (DFS) community is growing and the error rates for the forensic tools need to be continuously re-evaluated as the technology changes. This becomes more difficult in the case of mobile phone forensics, because they are proprietary. This paper discusses a database driven approach that could be used to store data about the mobile phone evidence acquisition testing process. This data can then be used to calculate tool error rates, which can be published and used to validate or invalidate the mobile phone acquisition tools.

Mobile Phone Usage

Mobile phones are widely used in the United States. In the first six months of 2006, the Cellular Telecommunication and Internet Association (CTIA) stated that there were 219.4 million U.S. wireless subscribers, and wireless communication has penetrated more than 72% of the total U.S. population. CTIA also explained that customers used 857 billion Minutes of Use (MOUs). Additionally, CTIA reported that 64.8 billion SMS messages were sent, an increase of 98.8% from the first six months of 2005 (“CTIA Quick Facts.” 2006).

Digital evidence is becoming important, where 80% of current court cases have some sort of digital evidence associated with them (Rogers, 2006, p.1). Summers (2003) explained “In the past five years, dozens of murderers have been convicted partly as a result of evidence about their mobile phones or those of their victims”. Mobile phones are becoming more than just simple phone devices. Numerous technologies are being integrated within them such as Bluetooth, digital cameras, Infrared, General Packet Radio Service (GPRS), E-mail and more.

Evidence needs to be acquired from mobile phones when needed in a forensically sound manner. In the realm of digital forensics, software tools have dominated the market in the acquisition of digital evidence from mobile phones. These tools have not been tested and have no published error rates. The only notable tool testing initiative for mobile phone forensics was performed by the National Institute of Science and Technology (NIST). This initiative is by no means complete, especially since they were only able to test a limited number of mobile phones, seventeen to be exact (Jansen, Wayne, Cilleros & Daniellou, 2005).

Forensic Tool Testing

Tool testing programs have been taken into consideration by various organizations. Tool testing is important from an Information Technology (IT) perspective to make sure that software and hardware operate as expected. The Institute of Electrical and Electronics Engineers (IEEE) has established standards since 1993 for tool testing. The International Organization for Standardization and the Electrotechnical Commission (ISO/IEC) then established the General Requirements for the Competence of Testing and Calibration Laboratories (ISO/IEC 17025) in 1999 ("General Testing Methodology.", 2001).

NIST's Computer Forensics Tool Testing (CFTT) program had the right intentions from a technical perspective, and NIST's ("General Testing Methodology.", 2001) states that the general requirements to test a tool are:

1. Establish categories of forensic requirements
2. Identify requirements for a specific category
3. Develop test assertions based on requirements
4. Develop test code for assertions
5. Identify relevant test cases
6. Develop testing procedures and method
7. Report test results

One can apply the aforementioned list to test a tool that is designed to work for a single, specific purpose, in an environment that is absolutely constant. However, in the case of mobile phones, numerous variables such as phone model, phone provider, cables used, even the fact that the mobile phone is on (data on a cell phone continuously changes when it is turned on) are all important factors that need to be properly documented.

Forensic Law at the Federal Level?

Ryan & Shpantzer (n.d.) explained that from 1923 until 1993, the admissibility of evidence was controlled by the Frye test, which states that expert scientific evidence is admissible only if the scientific community generally accepts it. In 1993, resulting from the *Daubert v Merrell Dow Pharmaceutical* (1993) case, the court adopted Rule 702 of the federal rules of evidence, which explained "If scientific, technical, or other specialized knowledge will assist the trier of fact to understand the evidence or to determine fact in issue, a witness qualified as an expert by knowledge, skill, experience, training, or education, may testify thereto in the form of an opinion or otherwise" (Ryan & Shpantzer, (n.d.), p. 2; Daubert v Merrell Dow Pharmaceutical, 1993).
The Daubert process has become synonymous with the admissibility of evidence at the federal level. The four major guidelines for the Daubert process that are used when assessing a procedure are:

1. Testing: Can and has the procedure been tested?
2. Error Rate: Is there a known error rate of the procedure?
3. Publication: Has the procedure been published and subject to peer review?
4. Acceptance: Is the procedure generally accepted in the relevant scientific community? (Carrier, 2002, p.3; "Digital Evidence," 2003)

When DFS scientists examine the above procedures, it is obvious that the science is deficient in the areas specified by the Daubert process. DFS' market driven nature has limited the amount of scientific research in tool testing. This has left scientists with an unclear understanding of the engineering behind these tools. The same applies to open source tools, because they are not documented properly and undergo constant change. This paper focuses on the first two guidelines of the Daubert process. The authors recognize the need for publishable error rates and the need for tool testing guidelines.

When the *Daubert v Merrell Dow Pharmaceutical* (1993) case is applied to digital evidence, it must satisfy two conditions 1) Evidence must be relevant (Federal Rules of Evidence 401, 2006) and 2) It must be "derived by the scientific method" and "supported by appropriate validation" (Ryan & Shpantzer, n.d.). One appropriate scientific consideration that is used for validation is the concept of repeatability.

Repeatability

The International Union of Pure and Applied Chemistry (IUPAC) defines repeatability as:

> The closeness of agreement between independent results obtained with the same method on identical test material, under the same conditions (same operator, same apparatus, same laboratory and after short intervals of time). The measure of repeatability is the standard deviation qualified with the term: 'repeatability' as repeatability standard deviation. In some contexts repeatability may be defined as the value below which the absolute difference between two single test results obtained under the above conditions, may be expected to lie with a specified probability ("Repeatability," 1997).

In order to establish repeatability, the conditions for its satisfaction as shown in the Guidelines for Evaluating and Expressing the Uncertainty of NIST Measurement Results by NIST's physics laboratory, include ("Guidelines for Evaluating.", 1994):

- The same measurement procedure
- The same observer
- The same measuring instrument, used under the same conditions
- The same location
- Repetition over a short period of time

It is important to consider the above definitions and requirements of repeatability when dealing with mobile phone tool testing. Tools should always acquire the same results, thus making a case for reproducibility. In traditional computer forensics, if one were to create a test case scenario for forensically analyzing a hard drive with a specific tool, it might not be so difficult to satisfy the above criteria, especially if one were using the same hard drive from the same manufacturer with the same configuration settings etc. This would be more difficult to achieve in mobile phone forensic tool testing since there are more variables that a test case would have to adhere to, such as mobile phone provider, locked down features etc. Other forensic disciplines have recognized the importance of error rates as well, such as DNA forensics which is considered to be more mature as a science (Saks & Koehler, 2005).

Testing in DNA Forensics

Forensic sciences have flaws. DNA forensics for example, has been widely accepted, yet even the results obtained from DNA forensics are not perfect. In a study by Saks & Koehler (2005), the following factors were shown to play a role in the wrongful conviction of 86 DNA exoneration cases:

1. 71% Eyewitness error
2. 63% Forensic science testing errors
3. 44% Police misconduct
4. 28% Prosecutorial misconduct
5. 27% False/misleading testimony by forensic scientists
6. 19% Dishonest informants
7. 19% Incompetent defense representation
8. 17% False testimony by lay witness
9. 17% False confessions

The two interesting statistics noted above are numbers 2 and 5. One can only imagine what the statistics would be like in the case of DFS, and it is our duty as Digital Forensic scientists to decrease those testing errors.

Systematic Database Driven Testing Methodology

The authors of this paper created a systematic database driven testing methodology for mobile phone tool testing. This will contribute to establishing repeatability estimates of the various tools that are used when acquiring digital evidence from mobile phones. With that comes a number of issues, mainly that mobile phones are proprietary. Therefore, a robust testing methodology for all mobile phones should take that into consideration.

Cellular Phones are Proprietary

New mobile phone models are released frequently by various corporations. It would be a difficult task to keep up with all these phone models and their various proprietary features. This poses a challenge in mobile phone forensic tool testing because a robust evidence acquisition system should be able to forensically acquire evidence from all mobile phone models even with the phone's software and hardware proprietary natures. Some of the proprietary mobile phone characteristics are outlined below. When dealing with mobile phone forensics, the following are important factors that should be recognized when performing a forensic acquisition:

1. Mobile phones have proprietary file systems.
2. Mobile phones have proprietary file transfer protocols.
3. Mobile phone providers lock down certain features of the device.
4. Different mobile phone providers might install different operating systems on the mobile phone device.
5. Cables used in the forensic acquisition of a mobile phone can be different.
6. The mobile phone device's clock changes data continuously on a the device.
7. Different mobile phones have different features.
8. A mobile phone being used is being provided a service through a carrier, and there are numerous carriers.
9. Applications can be installed on certain cellular phone models.

Process model for Cellular Phone Tool Testing

Based on the tool testing literature, a simplistic tool testing process model was developed. We envision the implementation of this process model as a programmatic database driven system. This process model is delineated in Figure 1.

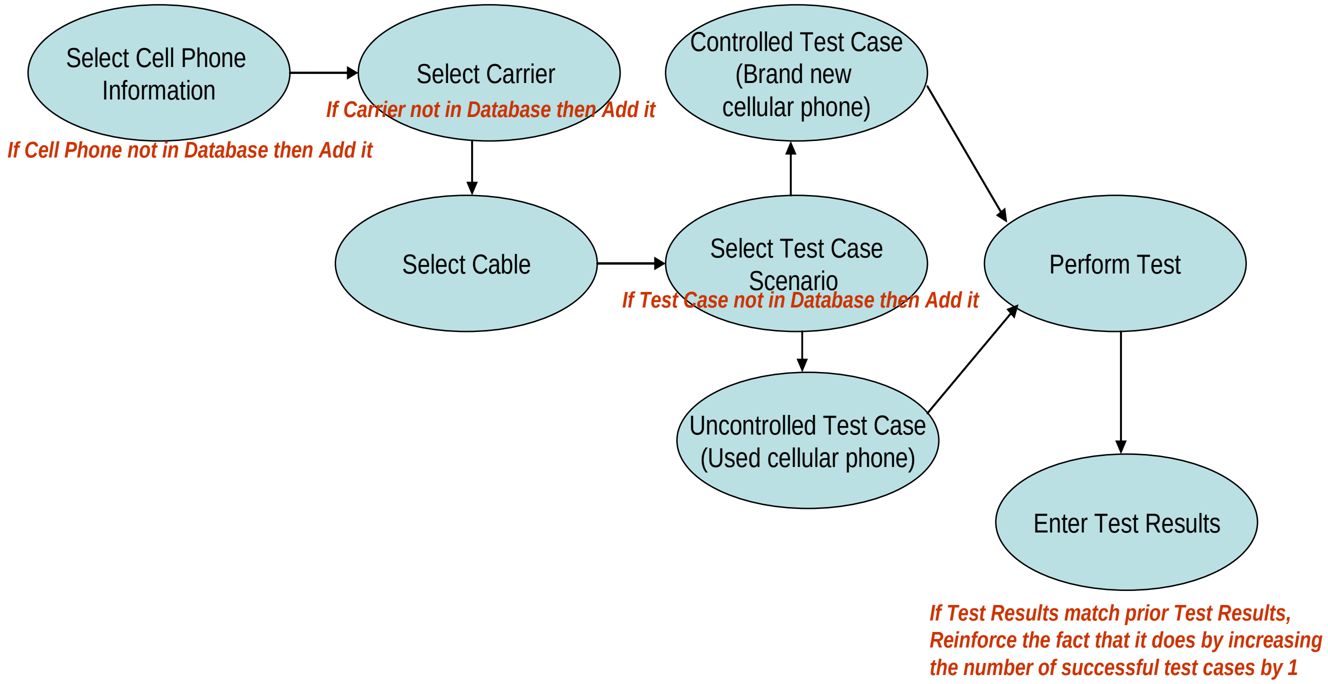
Select Cell Phone Information
Select Carrier
Controlled Test Case (Brand new cellular phone)
If Carrier not in Database then Add it
If Cell Phone not in Database then Add it
Select Cable
Select Test Case Scenario
Perform Test
If Test Case not in Database then Add it
Uncontrolled Test Case (Used cellular phone)
Enter Test Results
If Test Results match prior Test Results, Reinforce the fact that it does by increasing the number of successful test cases by 1

Figure 1- Cellular Phone Tool Test

As shown in Figure 1, the process model is simplistic in nature. The data that needs to be wrapped around that process model can be tedious. The above scenario is not necessarily new, as it adheres to the NIST tool testing methodology. The model simply takes the usual tool testing standards and tweaks them so that the process model is programmatically driven by a database system. Based on the process model and the proprietary nature of mobile phones, a relational database schema was developed to aid in illustrating the different data requirements for the forensic tool testing of mobile phones. Entity Relationship Diagrams (ERDs) are useful in representing data requirements. They adopt a more natural view that the real world consists of entities and relationships (Chen, 1976). The formulated database Entity Relationship Diagram is shown in Figure 2.

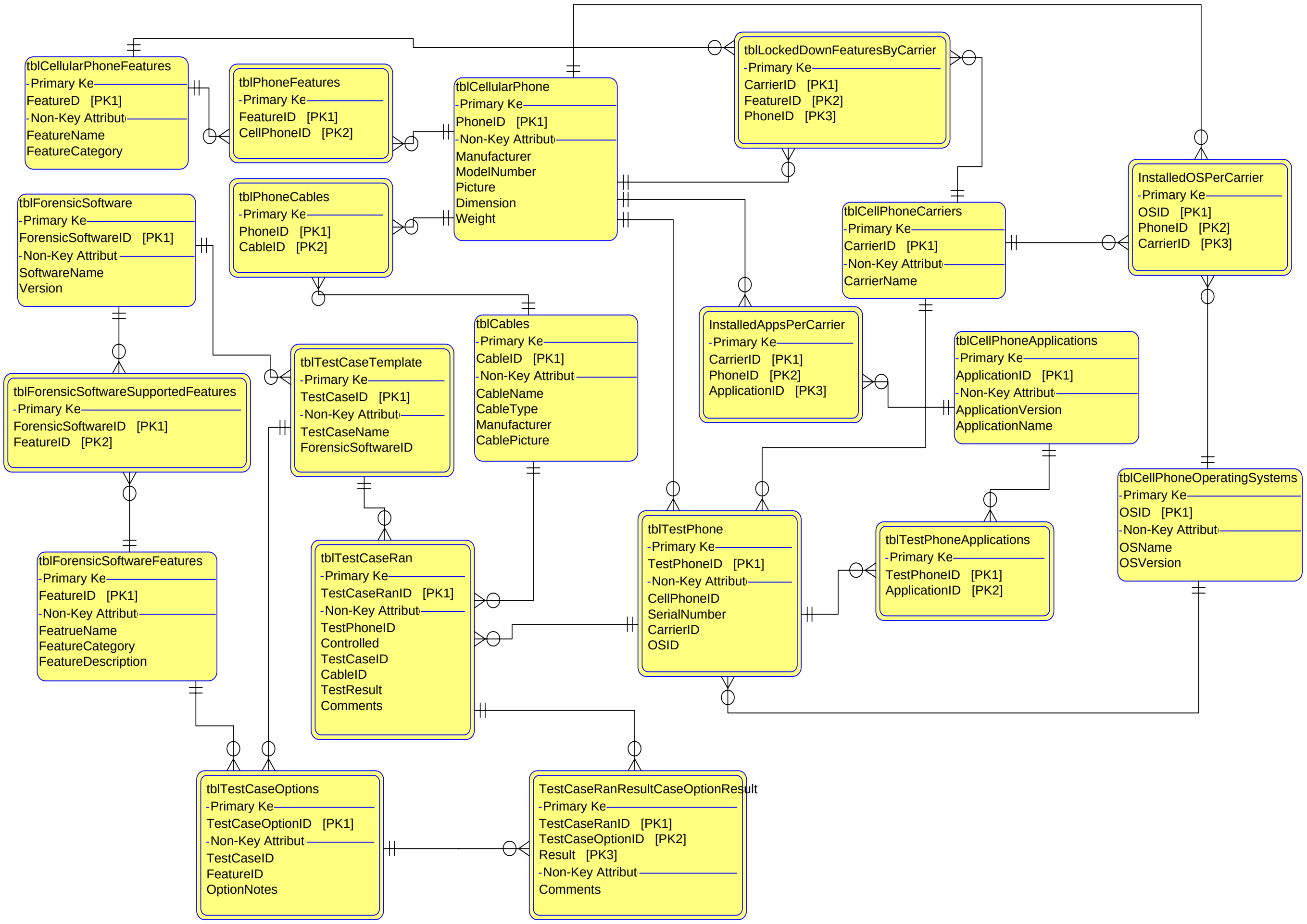

Figure 2- Proposed ERD

ERD Explanation

The ERD is merely the representation of the data that should be stored in the database The ERD above enables the users of a system to have access to the following data (assuming that the data has been entered into the database):

- The various mobile phone models
- The various features supported by each specific mobile phone
- The various test case scenarios
- The various mobile phone forensic acquisition tools
- The various features supported by each cellular phone acquisition tool
- Various data cables used in the forensic acquisition process
- The mobile phone models supported by each data cable
- The locked down features of a mobile phone that might be locked down by a carrier
- The various mobile phone carriers
- The operating system of a mobile phone installed by carrier for a specific phone model
- The various applications that might be installed by carrier for a specific mobile phone model
- Whether or not the overall test case passed or failed
- The features of the mobile phone tool test that passed or failed (For example, SMS acquisition passed, but call history failed)
- Various applications that could be installed on a mobile phone

The above data can help in obtaining specific information about the mobile phone being tested. It can also help in either validating or invalidating a forensic tool. For example, if a specific test case scenario were repeated, with the same conditions, yet with varying results, the tool would be deemed invalid. To know whether or not a tool is valid or invalid, error rates for the tools will have to be calculated.

Error Rate Types

Tool errors have been discussed in the literature (Carrier, 2002, 2003). The focus of this paper is not to create a topology of errors, but to illustrate a practical approach for the calculation of testing error rates. Using the database driven approach, one would be able to calculate the General Error Rate (GER) and the Feature Error Rate (FER) for every mobile phone test case scenario. These error rates can be calculated for every test case scenario as follows:

$$\text{General Error Rate (GER)} = \frac{\#\ \text{Unsuccessful forensic acquisition processes}}{\#\ \text{of Test Cases}}$$

$$\text{Feature Error Rate (FER)} = \frac{\#\ \text{Unsuccessful feature acquisition process}}{\#\ \text{of Feature acquisition processes}}$$

As a motivating example, assume that a mobile phone was analyzed and a test case scenario was performed. Different mobile phone devices that are of the same model were then used with the same test case scenario. For example, if 10 different mobile phones that are all Sony Ericsson K800i's were analyzed using the same test case scenario and the acquisition system worked only 4 times, then the GER = 6/10, which would equal 0.6 or 60%. Furthermore, if the software only fully worked 4 times, but it was able to acquire Short Message Service (SMS) messages, 7 times, then the FER_{SMS}= 3/10 which would be 0.3 or 30%.

Discussion of Using a Database Driven Approach

The proposed process model and database schema are advantageous for numerous reasons. Primarily, we as DFS scientists do not have access to any information about the tool error rates. This violates Daubert's second procedure as outlined by Carrier (2002). A database solution that logs all of the forensic examinations and tool testing procedures for the various mobile phones can help in the establishment of some type of calculated error rates based on the historical stability of these tool sets. This can also help in establishing a reliability measure for the various tool sets.

Once CF professionals become aware of the error rates of the forensic tools, it will help them become more objective with their decisions as to what tools perform the best functions. Furthermore, as mobile phone cases become more prominent in the courts, expert witnesses will be called upon. As an expert witness, it is vital to recognize the various error rates of the forensic tools to ensure that decisions are being made beyond a reasonable doubt.

Furthermore, having a standardized database driven approach for mobile phone forensics tool testing can help document the various testing scenarios that have been performed. The documentation of all the tests and their results create a historical repository that can be used for trend analysis, such as data mining. Statistical techniques can be applied to these historical results to help formulate predictions about the future forensic analysis of mobile phones by type, model, carrier etc. These results can further be used to ascertain the benefits inadequacies of the various mobile phone forensics tools.

Proposed Model Challenges

There are certainly some issues with using our proposed model. The first question that can be posed is how often are forensic professionals going to forensically acquire data from a mobile phone, and record the results using our anticipated database driven methodology? The authors believe there are two answers to this question. Primarily, we notice an increase of mobile phone ownership amongst college students. Therefore, scientists have access to a significant number of test cases at their disposal in learning institutions. Furthermore, the number of mobile phone models, and the number of mobile phones to be analyzed can come from that random sample of students through a research initiative. A database solution as such can be used by various investigators for either looking up past test results, or for adding their test case results. Of course, for that to occur, these investigators should be willing and able to provide us with that data. The solution the authors propose is not a definitive resolution to the error rate problem, but should spark some interest for further research in the area. Again, our purpose is to be able to find ways of attaining tool error rates, and furthermore, an estimate for the repeatability of the results that each mobile phone forensic acquisition tool provides.

Another issue with our proposed model is keeping the database up to date with the latest versions of the forensic acquisition tools and their capabilities. Some vendors release updates on a very frequent basis and their cooperation might be needed to keep the database current. A possible solution to this problem is to get the vendors involved in the process to continuously update the information in the database.

Conclusion

Mobile phone forensics is a novel field. When analyzing a mobile phone for forensic evidence, the process of doing so is different than the traditional computer forensics model. As forensic scientists, we should always be aware of the laws that deal with the admissibility of evidence, mainly the Daubert guidelines outlined by Carrier (2002) and in *Daubert v Merrell Dow Pharmaceutical* (1993). The first and second Daubert guidelines that deal with tool testing and error rates are the two major issues that this paper focused on. A database driven approach for the documentation of the mobile phone forensics procedures can ameliorate the process of documenting the testing methods. This will assist in acquiring results on the various test cases. These results can promote the calculation of tool testing error rates. This information will help DFS scientists validate or invalidate mobile phone forensics evidence acquisition tools and help expert witnesses make better decisions, beyond reasonable doubt about the evidence acquired from the mobile phones.

Future Work

The authors hope to implement this database driven system and perform a research study by acquiring mobile phones from students in an academic setting, then testing various acquisition tool sets on the market. These tests will allow the authors to publish tool testing error rates for the mobile phone acquisition tools.

About the Authors

Ibrahim Baggili is a doctoral student and graduate lecturer at Purdue University, West Lafayette, Indiana, in the department of Computer and Information Technology. His research interests include cyber forensics from a technical social and psychological perspectives and finding ways of improving the scientific validity of the field. His major current research initiative focuses on the effect of anonymity and integrity on cyber crime related activities. He can be reached at baggili@purdue.edu.

Richard Mislan is an Associate Professor of Computer and Information Technology at Purdue University, West Lafayette, Indiana. His research interests include Small Scale Digital Device (SSDD) forensics, unusual sources of digital evidence, and the application of artificial intelligence techniques for improving efficiency in cyber forensics. He can be reached at rmislan@purdue.edu.

Marcus Rogers is a Professor of Computer and Information Technology at Purdue University, West Lafayette, Indiana. His research interests include applied digital forensics, psychological crime scene analysis and information assurance. He can be reached at rogersmk@purdue.edu.

References

Ayres, R. Jansen, W. Cilleros, N. Daniellou, R. Cell phone forensic tools: an overview and analysis. Retrieved December, 11, 2006 from http://csrc.nist.gov/publications/nistir/nistir-7250.pdf

Carrier, B. (2002). Open source digital forensics tools: The legal argument. Retrieved December, 11, 2006 from http://www.digital-evidence.org/papers/opensrc_legal.pdf

Carrier, B. (2002). Defining digital forensics examination and analysis tools. Digital research workshop II. Retrieved February, 21, 2007 from at: http://www.dfrws.org/dfrws2002/papers/Papers/Brian_carrier.pdf

Carrier, B. (2003). Brian Carrier. Defining digital forensics examination and analysis tools using abstraction layers. International Journal of Digital Evidence, 1-4

Chen. P. (1976). The entity relationship model – toward a unified view of data. ACM Transactions on database systems, 1 -1, 9-36

Daubert v. Merrell Dow Pharmaceuticals. (1993). 209:579

Digital Evidence. (2003) Retrieved February 21, 2007, from http://www.golgotha.org.uk/academic/csmart01.html

Guidelines for evaluating and expressing the uncertainty of NIST measurement results. (n.d.). Retrieved September, 25, 2006 from http://physics.nist.gov/Pubs/guidelines/appd.1.html

Repeatability. (n.d.). Retrieved Sept 25, 2006, from http://www.iupac.org/goldbook/R05293.pdf

Rogers, M. (2006). DCSA: A Practical Approach to Digital Crime Scene Analysis. West Lafayette, Purdue University.

Ryan, D. Shpantzer, Gal. (n.d.). Legal Aspects of Digital Forensics. Retrieved December, 11, 2006, from http://www.danjryan.com/Legal Issues.doc

Saks, M. Koehler, J. The coming paradigm shift in forensic identification science. Science Magazine, 309, 892-895

Summers, C. (2003). Mobile phones - the new fingerprints. Retrieved February, 21, 2007 from http://news.bbc.co.uk/1/hi/uk/3303637.stm

Wireless quick facts. (2006). Retrieved September 23, 2006, from http://www.ctia.org/research_statistics/index.cfm/AID/10202