2011-04-21

Characterization of the Initial Flow Rate of Information During Reverse Engineering

Nicole Anderson
Brigham Young University - Provo

Characterization of the Initial Flow Rate of Information

During Reverse Engineering

Nicole Anderson

A thesis submitted to the faculty of
Brigham Young University
in partial fulfillment of the requirements for the degree of

Master of Science

Christopher A. Mattson, Chair
Brent L. Adams
John S. Lawson

Department of Mechanical Engineering

Brigham Young University

June 2011

ABSTRACT

Characterization of the Initial Flow Rate of Information
During Reverse Engineering

Nicole Anderson
Department of Mechanical Engineering, BYU
Master of Science

The future of companies that are founded on the development of new and innovative products is threatened when competitors reverse engineer and imitate the products. If the original developers could predict how long it would take a competitor to reverse engineer a product, it may be possible for them to delay, if not prevent, that competitor's entry into the market. Metrics and measures have been developed that can estimate the time it would take an individual to reverse engineer a product. The main purpose of these metrics and measures is to help designers determine how quickly a competitor could reverse engineer a product and develop and market a competing product. A critical parameter of these metrics is the flow rate of information (how quickly information can be extracted from a product), which is a parameter unique to each individual. This thesis seeks to establish a method for creating probability distributions that could be used to select a reasonable flow rate for an individual, by using data collected on the initial flow rate of multiple individuals.

Keywords: Reverse engineering, product development, flow rate of information, barriers to reverse engineering

ACKNOWLEDGMENTS

A special thanks to Dr. Mattson, Dr. Adams, and Dr. Lawson for all of their hard work, patients, and encouragement throughout the creation of this thesis. I'd also like to recognize my parents, A. Rick and Gail Anderson, for their immeasurable support during these long years of education; I wouldn't have made it without them.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

NOMENCLATURE

B	Barrier, or how difficult it is, to extract information about a product from the product itself
F	Rate at which information can be extracted from a product
K	Amount of information contained in a product
P	Power, or effort, exerted to extract information from a product
S	A measure of a product's ability to contain information
T	Estimated time to extract information K from a product

Subscripts, superscripts, and other indicators

$[\ ]_0$	indicates [ ] is evaluated at time t equal to zero
$[\ ]_i$	indicates [ ] is evaluated at time t equal to i
$[\ ]_j$	indicates [ ] is evaluated at time t equal to j

CHAPTER 1. INTRODUCTION

In the market place there are companies that are founded on the creation of new and innovative products or systems. The success of such companies is threatened when the products are reverse engineered and imitated by competitors. Since it is usually less expensive to imitate an already existing product, the competition can often sell their product at a lower price and take away part of the market from the originating company [1]. Shapiro [2] and Nelson and Winter [3] claim that the harder a product is to reverse engineer the less motivation there is for the competition to try and imitate it. The original developers would therefore benefit from a design methodology that could result in products that are difficult to reverse engineer [4, 5]. By making products that are more resistant to imitation the competition's entry into the market can be delayed, if not prevented. Thus the use of such a methodology would help the originating company maintain control of the market gained by the innovation [4].

One of the most important requirements of such a method is the ability to estimate the time it would take a competitor to reverse engineer the new product. By establishing a way to estimate how long it takes a competitor to extract a unit of information from a product it is possible to estimate how long it would take the competition to reverse engineer the product. With this knowledge, it is be possible to estimate the financial trade-offs between the cost of the additional features of the product, which make it more difficult to reverse engineer, and the rewards gained from the extended lag of the competitions market entry [4, 6].

Harston and Mattson [7] and Curtis and Harston and Mattson [8] have recently developed just such relations and metrics to predict the time it would take an individual to reverse engineer a product. One of the main parameters of these relations and metrics is the initial flow rate of information – the reciprocal of the time to gather a unit of information from a product, such as a feature dimension. The initial flow rate is approximated by the reciprocal of the time to make, and document, a measurement with minimal planning and validation. This thesis presents a method for creating probability distributions of initial flow rates of information types, and demonstrates the application of this method by creating a distribution for the initial flow rate of basic geometric information.

Harston and Mattson's metrics and methods model the time to reverse engineer a product with an exponential curve. Tests and studies indicate that different units of information in a product flow at different rates throughout the reverse engineering process. But Harston noticed that when these units of information are plotted from the fastest flowing to the slowest, they form a curve that can be closely approximated with an exponential function, provided the product is sufficiently complex. This exponential function can be determined so long as the slope, or flow rate of information, at at least one point is known [9].

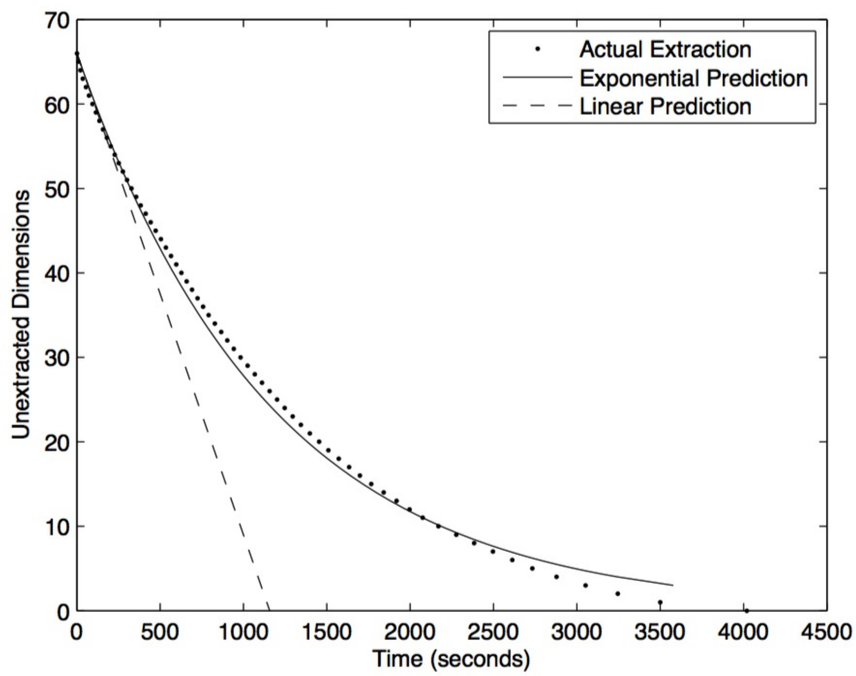


Figure 1.1: Actual and Estimated Times to Reverse Engineer a Sufficiently Complex Product

On the other end of the spectrum, it can be inferred that the time to reverse engineer products that are sufficiently simple could be estimated with a linear approximation. This approximation assumes that all the information in a product will be extracted at the same rate. The linear rate is the same as the initial rate of the exponential estimation. Therefore, the exponential estimation can be determined by using the slope of the linear estimation as the initial slope [9].

For the cases studied, these metrics and measures have estimated the time to reverse engineer a product to within an average error of 12.2%. But in each case an initial flow rate had to be determined specifically for the individual doing the reverse engineering. In other words, the initial flow rate of information is unique for each individual, as is the case with most human characteristics. This makes the parameter of initial flow rate potentially difficult to get into the hands of designers. One way to make this parameter more accessible is to determine a probability distribution that will show an average flow rate and range that most people who reverse engineer will fall within.

Current research indicates that a reasonable estimate of the total time to reverse engineer a product can be calculated if a value for the initial flow rate of information can be determined. But since this parameter is unique to each individual it is not readily accessible to designers. This thesis focuses on addressing this concern by presenting a method for determining a probability distribution of the initial flow rate of information types. An example of how to apply this method will also be given through the creation of a distribution of the initial flow rate of basic geometric information.

CHAPTER 2. TECHNICAL PRELIMINARIES; METRICS AND PARAMETERS FOR REVERSE ENGINEERING

This chapter presents a summary of the metrics for calculating the time to reverse engineer a product, previously developed by Harston and Mattson [7]. Since the work presented in this thesis is meant to support the work of Harston and Mattson it is important to understand the basic relationships behind their metrics as a foundation for the developments presented in this thesis.

The metrics and parameters presented in this section have already been developed by Harston and Mattson, so the full development will not be shown here [7]. It should be noted however that these metrics were developed from the electrical relationship known as Ohm's law, where the flow of information from a product corresponds with the flow of current out of a capacitor through a resistor.

Just as a circuit has a certain level of capacitance, or ability to store a charge, a product will have a certain ability to store information. The ability of a product to store information is characterized by

$$S_j = \frac{K_j F_j}{P} \tag{2.1}$$

where F_j is the flow rate of information, or the rate at which information can be extracted from a product, at time t equal to j; P is the power, or effort per time, being exerted to extract information; and K_j is the amount of information stored in a product at time t equal to j. The value of P is constrained by

$$0 < P \leq 1 \tag{2.2}$$

where zero represents no effort being exerted to reverse engineer the product and one indicates that the greatest possible effort is being exerted. For this thesis we assume that P is always 1. The value of K_j is constrained by

$$0 < K_j < K_0 \tag{2.3}$$

where K_0 is the amount of information stored in the product at time t equals zero. The flow of this information out of a product will be hindered by any barriers to reverse engineering that are in the product; much like a resistor hinders the flow of current out of a capacitor. A product's barrier to reverse engineering is given by

$$B_j = \frac{P}{F_j^2} \tag{2.4}$$

The time to extract the information in a product, or the time to reverse engineer a product, is given by the following exponential decay relationship

$$T_i = -B_j S_j \ln\left(\frac{K_i}{K_0}\right) \tag{2.5}$$

where T_i is the estimated time to extract all but K_i – the amount of information remaining in the product at time t equals i – units of information from the product. While this relationship may not directly appear to rely on the flow rate of information, by substituting equations (2.4) and (2.1) into equation (2.5)

$$T_i = -\left(\frac{P}{F_j^2}\right)\left(\frac{K_j F_j}{P}\right)\ln\left(\frac{K_i}{K_0}\right) \tag{2.6a}$$

$$T_i = -\frac{K_j}{F_j}\ln\left(\frac{K_i}{K_0}\right) \tag{2.6b}$$

the reliance becomes readily apparent. In words, equation (2.6b) states that the time to reverse engineer a product is inversely proportional to the rate at which information can be extracted from that product. This relationship makes sense; the faster information can be extracted from a product, the less time it will take to reverse engineer it.

Since the K values are determined from the design data, F is the only unknown in Equation (2.6b). The flow rate of information is determined from the individual reverse engineering the product. Because of it's individualistic nature, determining a value for F can be challenging and time consuming, which reduces the effectiveness of using the metrics and measures. This thesis focuses on determining a method that will allow designers to determine values for the flow rate of information simply and quickly, thus increasing the effectiveness of the metrics and measures.

CHAPTER 3. A METHOD FOR CREATING DISTRIBUTIONS OF INITIAL FLOW RATES OF INFORMATION

The flow rate of information is a parameter unique to each individual, just like the speed with which a person runs a mile is unique to each individual. While the time to run a mile is unique, there is a spread, or distribution, in which people tend to fall. This chapter presents the theory that distributions of initial flow rates of information types also exist, and that those distributions can be quantified easily without the need to fully reverse engineer complex products. The general process for quantifying these distributions includes four main steps;

1. Determine the most elementary unit of the information type for which the distribution will be created; this should have the fastest flow rate.
2. Find, or create, a test speciman that is simple and only consists of the elementary unit.
3. Measure the time it takes people to extract the information from the test specimen and divide the number of units in the test specimen by the time it took to extract the information; these are the initial flow rates for the information type.
4. Perform statistical analysis and create a distribution of the initial flow rates.

The next section will discuss these steps further and will introduce a specific application that will be used throughout the rest of the thesis.

3.1 Determining the Complex Estimation with the Simple Estimation

Harston and Mattson noticed that only products that are "sufficiently complex" match the exponential estimate; which is calculated using Equation (2.6b). It should be noted that most products out in the market can be considered sufficiently complex, like the tape case seen Figure 3.1. Products such as washers, cubes, and the like are examples of products that are not sufficiently complex. Of course, if only "sufficiently complex" products match the exponential estimation, it seems reasonable to infer that only products that are "sufficiently simple" will match the linear estimation [7, 9]. A sufficiently simple product is one where every unit of information is extracted at the same rate. Products that are not sufficiently complex or simple will fall somewhere between the linear and exponential estimations, as seen in Figure 3.2.

Figure 3.1: A Tape Case is a Sufficiently Complex Product

When plotted on the same graph it is easy to see that the slope of the linear estimation matches the initial slope of the exponential estimation. Therefore, if the time to fully reverse engineer a sufficiently simple product could be captured, the ratio of amount of information to extraction time would determine the initial flow rate of the exponential estimate for the sufficiently complex products. Hence the reason the first step to creating a distribution of initial flow rates of an information type is to determine what, within the information type of interest, the most elementary unit of information is. A sufficiently simple product will be made up solely of this elementary unit of information. The initial flow rate of the information type, for an individual, is calculated by dividing the amount of information in the product by the time it takes the individual to extract that information. With enough initial flow rates (≥ 15) a distribution can be created and characterized.

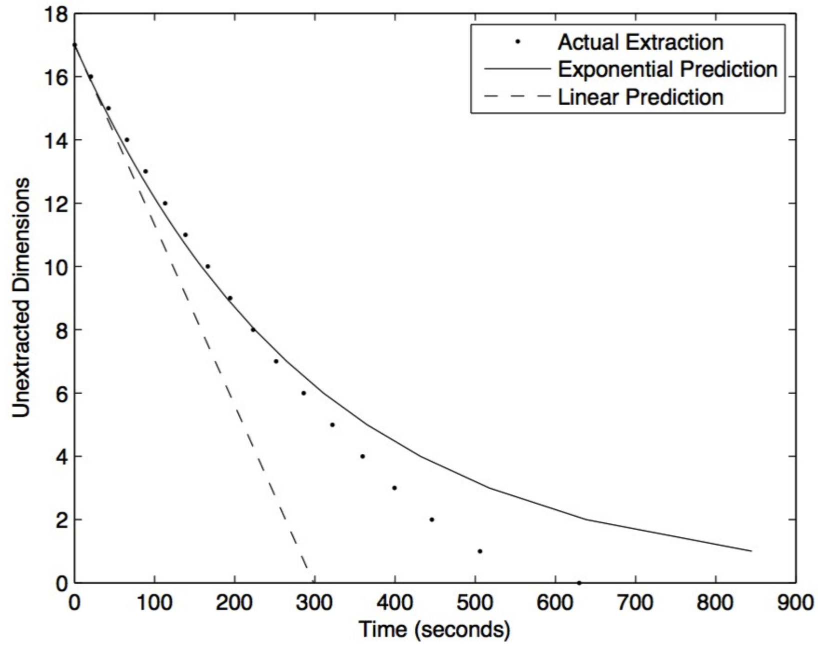


Figure 3.2: Reverse Engineering Time Estimations for a Product that is not Sufficiently Simple or Complex

For this thesis we will look at applying the process for creating an initial flow rate distribution to basic geometric information. We defined the set of basic geometric information to contain only linear geometric dimensions. The most elementary unit of information was determined to be simple length measurements that can be measured with a pair of digital calipers. Therefore the sufficiently simple product was a parallel piped (cuboid), Figure 3.3, that was approximately 1.996 *X* 1.047 *X* 0.186 inches.

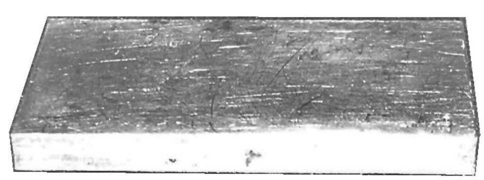


Figure 3.3: A Cuboid, that is neither too Big or Small, is a Sufficiently Simple Product for Basic Geometric Information

The following section will discuss the data collection for the distribution of initial flow rates of basic geometric information.

3.2 Data Collection

For the example in this thesis the target population was defined to be those who reverse engineer with the understanding that a part, or all, of the product will be replicated. Three main characteristics of this population are:

1. A basic understanding of the tools (calipers, for the example in this thesis)

2. An ability to identify the information that defines what a product is or does

3. An ability to determine how to collect the defining information of a product

We determined that these characteristics are most likely to be found in individuals who are in, or going into, a technical field (mechanical and manufacturing engineering in particular). Therefore we targeted those in mechanical or manufacturing engineering fields, or in an academic program for mechanical or manufacturing engineering. The data was divided into three groups:

1. Those in an academic mechanical or manufacturing engineering program who have completed an introduction to CAD class, with limited to no experience in a technical field

2. Those in an academic mechanical or manufacturing engineering masters program, with limited to no experience in a technical field

3. Those experienced in a technical field

Some simple code written, using a computer program called MatLab, that asked individuals to draw a picture/drawing of the cuboid, measure the three dimensions, and then record those measurements on the drawing they had created. The program captured the time it took each person to draw, and measure and record the dimensions of the cuboid. Using this test, we were able to collect about 30 data points from each group, where a data point consisted of 1 individual's average flow rate. Approximately 1/3 of the data points came from individuals at 2 technical companies. In the next chapter we explain the analysis of this data.

CHAPTER 4. STATISTICAL ANALYSIS OF THE INITIAL FLOW RATE OF BASIC GEOMETRIC INFORMATION DATA

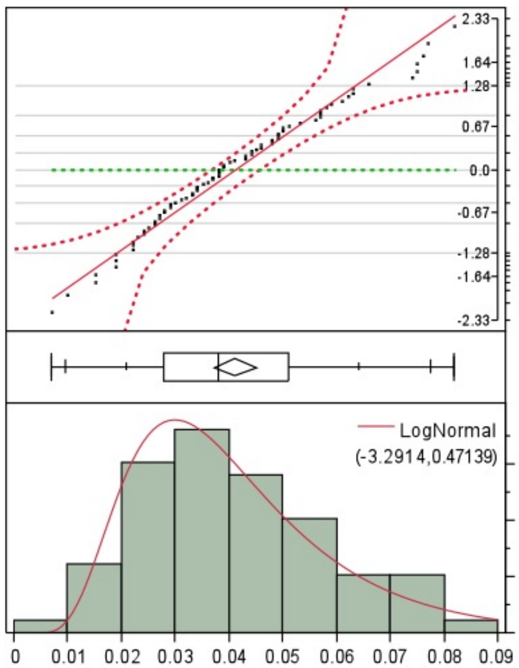
Four distributions were created from the data; a master distribution, which included all the data points; and then three distributions with just the data points from the individual groups, these groups will be refer to collectively as the subgroups. As seen in Figure 4.1, three plots were created for each of the four data groups; a normal quantile plot, a quantile box and whisker plot, and a data distribution plot with a fitted lognormal distribution overlaid. These plots provide a graphical indication of the spread and behavior of the data.

The normal quantile plots were created to determine how well one of the normal distribution describes the distribution of the data. In order to reject the null hypothesis – that the data comes from one of the normal distributions – the data points would fail to fall along the straight line on the quantile plot. If the data does fall on along the straight line the data will not have provided enough evidence to reject the null hypothesis, and we can reasonably conclude that one of the normal distributions may be used to represent the data.

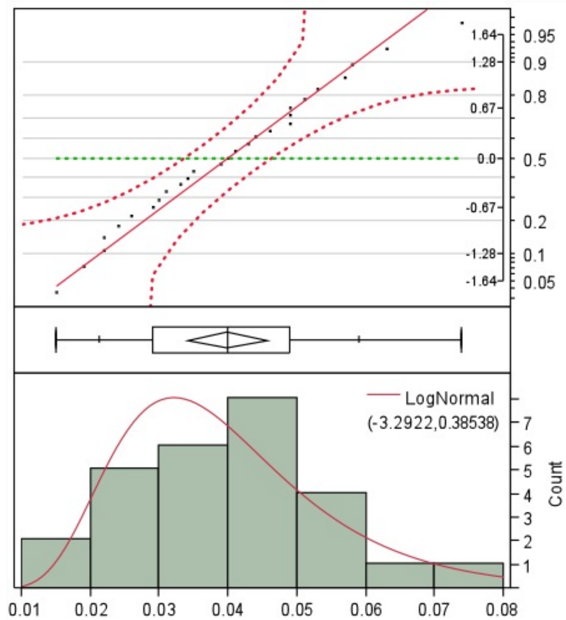
Since it is physically impossible to have a negative flow rate, a goodness-of-fit test was performed on a fitted lognormal distribution to determine the likelihood that the data would come from from a lognormal distribution. In order to reject the null hypothesis – that the data comes from a lognormal distribution – the the p value result for the goodness-of-fit test would need to be less than, or equal to, 0.05. If the p values are greater than 0.05 there will not be enough evidence to reject the null hypothesis, and we can reasonably conclude that the data comes from a lognormal distribution.

Quantile Box and Whisker plots were created for the four groups to provide a graphical image of the data spread. These plots consist of three main categories of markings; box, diamond, and whiskers. The center line of the box indicates the median, or 50% mark, of the data. For a lognormal distribution this value is calculated by exponentiating the μ parameter of a lognormal distribution. The ends of the box represent the 25% mark and the 75% mark of the data. In the middle of the box is a diamond. The vertical center of the diamond represents the mean, or average, of the data. This value is not utilized in the analysis of the data considered in this thesis because of the use of the lognormal distribution.

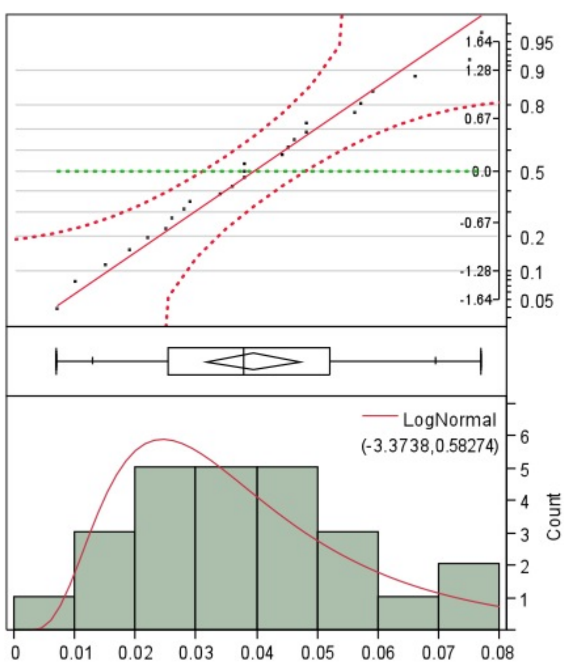
By comparing the median of the three subgroups with the master median, it is possible to determine if the amount of industry experience an individual has influences that individual's initial flow rate. The null hypothesis is that industry experience does not affect an individual's initial flow rate. In order to reject this null hypothesis, at least one of the subgroups would need to have its median fall at a statistically significant distance from the master median. If this should happen, it would be reasonable to assume that the individuals represented by that subgroup will have flow rates that come from a different distribution. Therefore those individuals will tend to have flow rates that compare to the average population in the same way that the subgroup's median compares to the master distribution's median. For example, assume that data group 3 (containing individuals who are experienced in industry) has a median that lies so far into the master distribution's faster region that the difference between the group's median and the master's median is statistically significant. In this case we would expect that individuals who fit into group 3 to generally have faster flow rates than the average population. However, if none of the subgroups have a median that falls at a statistically significant distance from the master median, there will not be enough evidence to reject the null hypothesis. In this case it will be reasonable to claim that an individuals initial flow rate does not depend on his/her amount of industry experience.



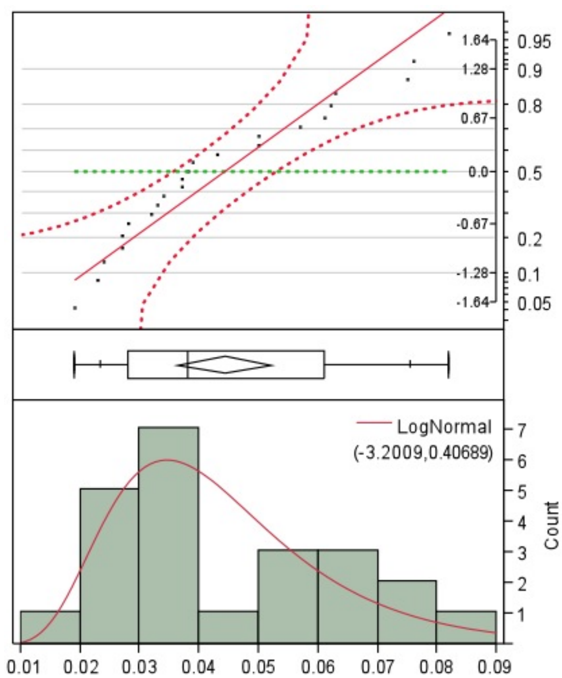
(a)



(b)



(c)



(d)

Figure 4.1: Analysis plots: (a) Master Plots; (b) Group 1 Plots; (c) Group 2 Plots; (d) Group 3 Plots.
Top) Normal Quantile Plot; Middle) Quantile Box & Whisker Plot; Bottom) Data Distribution Plot with Fitted Log-normal Distribution Overlaid

CHAPTER 5. DISCUSSION OF RESULTS

As was mentioned in Chapter 4, the lognormal was selected as the distribution to represent the data because it illustrates distributions which cannot have negative values, and it is impossible to have a negative flow rate of information. However, it should be pointed out that it is not uncommon for the parameters of a lognormal plot to be negative. While, with a normal distribution, μ and σ would typically represent, respectively, the mean and standard deviation of the actual flow rate values, this is not the case for a lognormal distribution. The μ and σ for a lognormal plot represent, respectively, the mean and standard deviation of the natural log of the flow rate values. Therefore, when we use the standard statical calculations to determine a value, we must perform the opposite operation of the natural log, the exponential, in order to determine the actual value we want. For example, in order to determine the actual value of a flow rate that is 3σ from the mean we would calculate $\mu + 3\sigma$, as usual, but then we take the exponential of that value, $e^{\mu+3\sigma}$.

Table 5.1 shows the lognormal parameters, along with the exponentiated values of those parameters, and the p values of the goodness-of-fit test of the entire data set and the three subsets. By looking at the p values it is clear that there is not enough evidence to reject the null hypothesis that the data comes from a lognormal distribution. Also, there is no statistical difference between the medians of the groups. Therefore it is reasonable to claim that experience in industry does not affect an individuals initial flow rate for basic geometric information.

To ensure that these results made sense, we timed an individual reverse engineering the basic geometric information of an Apple® keyboard. From past evidence we expected the time to fall in the upper range of the distribution. A plot was created which estimated the time it would take to reverse engineer the keyboard for an individual with the median, or e^{μ}, flow rate of basic geometric information, an individual with a flow rate of basic geometric information equal to $e^{\mu+3\sigma}$ (which is the equivalent to $\mu + 3\sigma$ for a regular normal distribution), and the actual time it took the individual to reverse engineer the keyboard. As expected, the actual time fell between

Table 5.1: Data Group Comparison

Group:	All data points	Data points of undergraduate students	Data points of graduate students	Data points of those experienced in industry
μ	-3.291	-3.292	-3.374	-3.201
Median	0.037	0.037	0.034	0.041
σ	0.471	0.385	0.583	0.407
e^{σ}	1.602	1.470	1.791	1.502
p value	0.0687	0.1051	0.1348	0.1076

the two estimated times. This indicates that the presented distribution of initial flow rates of basic geometric information is a reasonable representation for people who reverse engineer products.

Figure 5.1: Keyboard Before Disassembly

3σ was selected because of its significance and heavy use in industry to determine if a system is out of control or not. When creating a control chart there are three critical lines: the upper control limit, the median/mean, and the lower control limit. The upper and lower control limits are drawn at $\mu+3\sigma$ and $\mu-3\sigma$, respectively. Any one point that falls outside of either of the 3σ lines is considered "out of control" and is likely influenced by something other than natural variation [10, 11]. Therefore it is reasonable to assume that about 99.7% of individuals will have a flow rate of information between the 3σ marks.

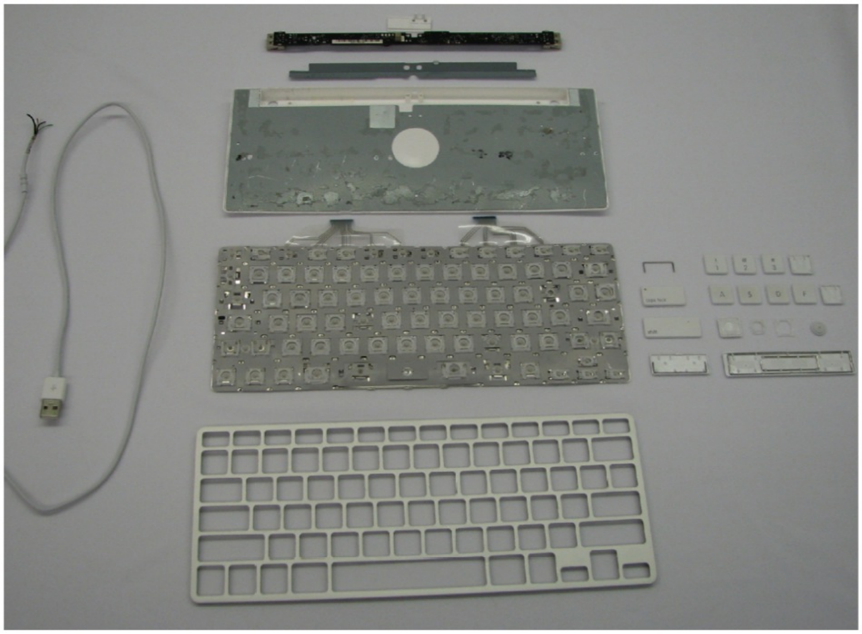


Figure 5.2: Keyboard After Disassembly

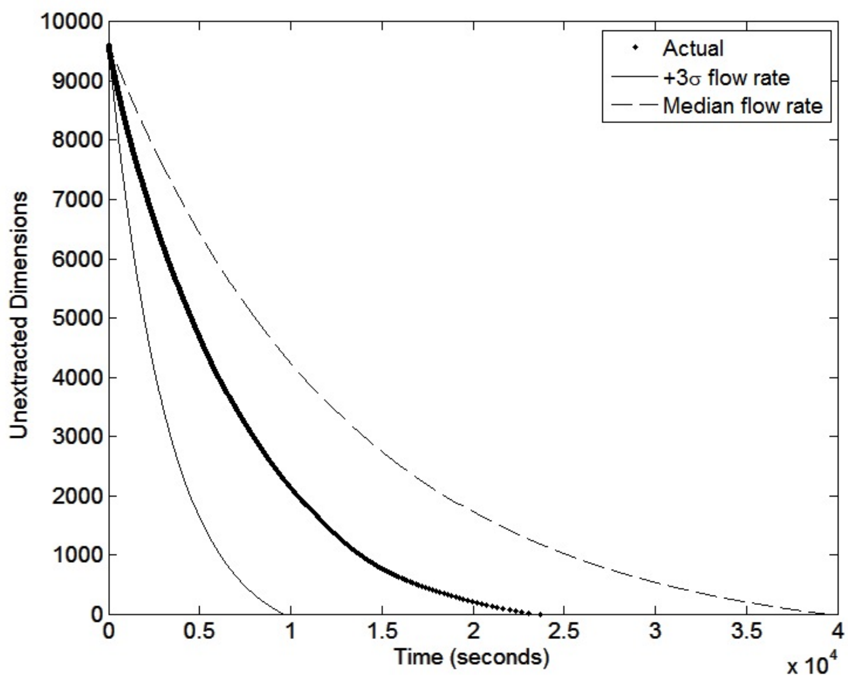


Figure 5.3: Actual and Estimated Reverse Engineering Times for an Apple® Keyboard

CHAPTER 6. CONCLUDING REMARKS

Metrics and measures have been developed that can estimate the time it would take an individual to reverse engineer a product, known as the Barriers to Reverse Engineering (BRE) metrics and measures. These metrics and measures have the potential to help designers estimate how long it would take a competitor to reverse engineer a product and then develop and market a competing product. Unfortunately such calculations require knowing the value of a parameter that is unique to each individual; the flow rate of information. The individualistic nature of this parameter makes it difficult for designers to use the BRE metrics and measures.

This thesis focused on addressing this concern by determining a process by which a distribution for the initial flow rate of a type of information can be created. The process is based on the fact that the exponential and linear reverse engineering time estimations have have the same slope, or flow rate of information, at the time t equal to zero. Therefore the exponential estimation can be determined with the slope of the linear estimation. To illustrate this process, a distribution was determined for basic geometric information. To help ensure that the distribution made sense we use the distribution to estimate a range between which we expected an individual to fall while reverse engineering an Apple® keyboard and then actually timed the individual as they reverse engineered the keyboard. The actual time fell well within the range we had predicted and so it seems reasonable to conclude the distribution for basic geometric information is likely to fit most people who reverse engineer things.

By establishing a process by which the individualistic nature of the flow rate of information can be characterized, we have made the BRE metrics and measures more usable. Designers may now more easily use the BRE metrics and measures to estimate the time it will take a competitor to reverse engineer a product and develop and market a competing product. Being able to make such predictions has the potential to help protect companies that depend on the development of new and innovative to give them a competitive advantage.

REFERENCES

[1] Schnaars, S. P., 1994. *Managing Imitation Strategies: How Later Entrants Seize Markets From Pioneers*. The Free Press, New York, New York. 1

[2] Shapiro, C., 1985. “Patent Licensing and R&D Rivalry.” *American Economic Review,* **75**(2), May, pp. 25–30. 1

[3] Nelson, R. R., and Winter, S. G., 1982. *An Evolutionary Theory of Economic Change*. Harvard University Press, Massachusetts. 1

[4] Knight, D. C., Mattson, C. A., and Adams, B. L., 2009. “Maximizing Return on Investment by Constructing Optimal Barriers Against Competitors’ Market Entry.” In *5th AIAA MDO Specialists Conference*, AIAA, pp. 2009–2190. 1

[5] Gruca, T. S., and Sudharshan, D., 1995. “A Framework for Entry Deterrence Strategy: The Competitive Environment, Choices, and Consequences.” *The Journal of Marketing,* **59**(3), July, pp. 44–55. 1

[6] Knight, D. C., 2011. “Metrics and parameters for evaluating the return on investment of implementing barriers to reverse engineering and imitating.” Masters of science, Brigham Young University, Provo, UT, April. 1

[7] Harston, S. P., and Mattson, C. A., 2009. “Metrics for Evaluating and Optimizing the Barrier and Time to Reverse Engineer a Product.” In *2009 Design Engineering Technical Conference, and Computers and Information in Engineering Conference*, no. DETC2009-86781. 1, 5, 8

[8] Curtis, S. K., Harston, S. P., and Mattson, C. A., 2009. “A Generic Formulaic Characterization of the Time to Reverse Engineer the Tolerances of a Product.” In *ASME 2009 International Mechanical Engineering Congress and Exposition*, Vol. 4, ASME, pp. 275–285. 1

[9] Harston, S. P., 2009. “A Methodology for Designing Product Components with Built-in Barriers to Reverse Engineering.” Masters of science, Brigham Young University, Mechanical Engineering, Provo, UT. 2, 8

[10] Balestracci, D., 2006. “Data ’sanity’: Statistics and reality.” *Quality in Primary Care,* **14**(1), February, pp. 49–53. 16

[11] Balestracci, D., 2006. “Statistics and reality: part 2.” *Quality in Primary Care,* **14**(2), June, pp. 111–119. 16

APPENDIX A. MATLAB CODE FOR COLLECTING DATA ON THE FLOW RATE OF BASIC GEOMETRIC INFORMATION

The following is the MatLab code that was used to collect the data which was used to create the distribution for the initial flow rate of basic geometric information. This code can also be used to determine an individual's average initial flow rate, thus removing the need to use the distribution to determine an appropriate initial flow rate as the exact value can now be used instead. We are providing this code as another way for designers to determine values for F and thus increase the effectiveness of the BRE metrics and measures.

```
% Display instructions to the screen to help orient the user
% Explain the test
display('TEST INSTRUCTIONS')
display(' ')
display('There are two sections to this test, and you will need to press ENTER')
display('to begin and end each section.')
display(' ')
display('In Section 1 you will create a sketch of an aluminum block.')
display('Do NOT draw dimension lines during this section of the test.')
display(' ')
display('In Section 2 you will measure the three dimensions of the block')
display('and record them on the sketch you created in Section 1.')
display('You may draw dimension lines during this section.')
display(' ')
display('Make all measurements as if it were your job to replicate the block.')
display('You will be asked to report your measurements at the end of the test.')
```

```
% Prompt the user to go forward with the test when they are ready
display(‘ ’)
display(‘Press ENTER when you are ready to begin the first section.’)
pause;

clc

%%%%%%%%%%%%%
% Read from the right counter file to keep track of the runs
Counter = [‘Counter.txt’];
testrun = dlmread(Counter);

% Increment the Counter File
newTestRun = testrun + 1;
dlmwrite(Counter,newTestRun);
%%%%%%%%%%%%%

% Instruct the user to draw the picture on which they will record
% their measurements
display(‘SECTION 1: Create a sketch of the block.’)
display(‘ ’)
display(’Press ENTER when you”re ready to begin drawing, and again when you”re done.’)
display(‘ ’)
display(’REMEMBER: Do NOT draw any dimension lines yet, just the block.’)
display(‘ ’)
pause;

% Time the drawing clock
display(‘ ’)
display(‘Drawing...’)
```

```
DrawStart = tic;
pause;
DrawTime = toc(DrawStart);
clc;

% Start the test
display('SECTION 2: Measure and record the three dimensions of the block.')
display(' ')
display('Press ENTER when you''re ready to begin, and again when you''re done.')
display(' ')
pause;

% Time the dimension extraction
display(' ')
display('Measuring...')
Begin = tic;
pause;
Time = toc(Begin);
clc;

%%%%%%%%%%%%%

% Ask the user for their measurements
clc;
display('What was the measurement for the long side?')
try
longSide = input('\n');
while isempty(longSide),
longSide = input('Please enter the value you measured for the long side: ');
end
```

```
catch ME
display('Do not type any letters when reporting your measurements.')
longSide = input('Please re-enter your measurement for the long side: ');
end

display(' ')
display('What was the measurement for the thickness of the block?')
try
thick = input('\n');
while isempty(thick),
thick = input('Please enter the value you measured for the thickness: ');
end
catch ME
display('Do not type any letters when reporting your measurements.')
thick = input('Please re-enter your measurement for the thickness: ');
end

display(' ')
display('What was the measurement for the other side?')
try
width = input('\n');
while isempty(width),
width = input('Please enter the value you measured for the other side: ');
end
catch ME
display('Do not type any letters when reporting your measurements.')
width = input('Please re-enter your measurement for the other side: ');
end
dims = [longSide width thick];
```

```
% Ask the user about their industrial experience
display(‘Which of the following best describes you? (Enter 1, 2, or 3)’)
display(‘ ’)
display(’1) In an acedmic mechanical or manufacturing engineering program, have completed an’)
display(‘ introductory class on CAD, and have limited or no experience in a technical field’)
display(‘ ’)
display(‘2) In an acedmic mechanical or manufacturing engineering masters program,’)
display(‘ with limited or no experience in a technical field’)
display(‘ ’)
display(‘3) Experienced in a technical field’)
try
experience = input(‘\n’);
test1 = experience == 1;
test2 = experience == 2;
test3 = experience == 3;
test = test1 + test2 + test3;
while test  = 1,
experience = input(‘Please type the number next to the option which best describes you. ’);
test1 = experience == 1;
test2 = experience == 2;
test3 = experience == 3;
test = test1 + test2 + test3;
end
catch ME
experience = 0;
end

% Identify the proper file to write the data into based on experience
analysisFile = [‘CuboidAnalysis’ num2str(experience) ‘.txt’];
```

```
% Calculate how accurate each measurement was
longSideAcu = (abs(1.996 - longSide)/1.996)*100;
widthAcu = (abs(1.047 - width)/1.047)*100;
thickAcu = (abs(0.186 - thick)/0.186)*100;
aveAcu = (longSideAcu + widthAcu + thickAcu)/3;

% Tell the user they are finished
clc;
display(‘Thanks! You”re finished!’)
display(‘ ’)
showMe = sprintf(‘You”re measurements had an average error of %0.3f%%’,aveAcu);
disp(showMe)
display(’Great job!’)

%Calculate the flow rate as the amount of information extracted
%divided by the time.
K = 3;
F = K/Time;

% Write the results to a file
fid = fopen(‘CuboidResults.txt’,‘a’);
fprintf(fid,‘————————————————————————————\nTest Run: %u\n’,testrun);
fprintf(fid,‘Experience: %u\n’,experience);
fprintf(fid,‘\nYear: %u Month: %u Day: %u\nTime: %u:%u and %u sec\n’,clock);
fprintf(fid,‘\nTotal time to create sketch: %4.4f seconds\n’,DrawTime);
fprintf(fid,‘Total time to extract %u dimensions: %4.4f seconds\n’,K,Time);
fprintf(fid,‘Length: %0.3f\tWidth: %0.3f\tThickness: %0.3f’,dims);
fprintf(fid,‘\nFlow rate: %0.3f\n’,F);
fclose(fid);
```

```
% Create data files for automatic analyzation of the data
% General file
fid = fopen('CuboidAnalysis.txt','a');
fprintf(fid,'%u\t',testrun);
fprintf(fid,'%u\t',experience);
fprintf(fid,'%4.4f\t',DrawTime);
fprintf(fid,'%4.4f\t',Time);
fprintf(fid,'%4.3f\t',F);
fprintf(fid,'%0.3f\t%0.3f\t%0.3f\t',dims);
fprintf(fid,'%0.3f\n',aveAcu);
fclose(fid);
% Experience specific file
fid = fopen(analysisFile,'a');
fprintf(fid,'%u\t',testrun);
fprintf(fid,'%u\t',experience);
fprintf(fid,'%4.4f\t',DrawTime);
fprintf(fid,'%4.4f\t',Time);
fprintf(fid,'%4.3f\t',F);
fprintf(fid,'%0.3f\t%0.3f\t%0.3f\t',dims);
fprintf(fid,'%0.3f\n',aveAcu);
fclose(fid);
```