THE GEOSPATIAL FOOTPRINT OF MORTGAGE FRAUD: A CASE STUDY IN ALACHUA COUNTY

By

AUSTIN EVANS

A THESIS PRESENTED TO THE GRADUATE SCHOOL OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF ARTS

UNIVERSITY OF FLORIDA

2011
To my family, who has been so supportive of my pursuit — I never would have been able to reach this goal without your help and encouragement.
ACKNOWLEDGMENTS

First and foremost, I would like to acknowledge the support and help of my advisor, Dr. Grant I Thrall. Dr. Thrall has been instrumental in providing ideas, direction, and feedback throughout the writing of this paper. It has been an honor to work with a true leader in the field.

I would also like to acknowledge the support of my committee members, Dr. David Ling and Dr. Tim Fik. They provided me with the background knowledge I needed to make this paper possible, as well as key ideas and guidance throughout this process.

The support of Chris Davis at the Herald Tribune must be acknowledged as well. I greatly appreciate him sharing the data and techniques that he used in his investigation, and his help implementing them into my research.

I must also express my appreciation for the efforts of Desiree Price, who always had the answers to my questions and kept me on track. I would never have made it through all of the administrative issues and requirements without her help and well timed reminders.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>SECTION</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACKNOWLEDGMENTS</td>
<td>4</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>6</td>
</tr>
<tr>
<td>ABSTRACT</td>
<td>7</td>
</tr>
<tr>
<td>CHAPTER</td>
<td></td>
</tr>
<tr>
<td>1 MORTGAGE FRAUD AND GEOGRAPHY</td>
<td>9</td>
</tr>
<tr>
<td>Summary</td>
<td>9</td>
</tr>
<tr>
<td>Background</td>
<td>12</td>
</tr>
<tr>
<td>2 ANALYSIS</td>
<td>18</td>
</tr>
<tr>
<td>Data Filtering</td>
<td>18</td>
</tr>
<tr>
<td>Making the Data Geospatial</td>
<td>20</td>
</tr>
<tr>
<td>Grid Cell Approach</td>
<td>20</td>
</tr>
<tr>
<td>Dot Density Approach</td>
<td>22</td>
</tr>
<tr>
<td>3 RESULTS</td>
<td>25</td>
</tr>
<tr>
<td>Gaussian Smoothing</td>
<td>25</td>
</tr>
<tr>
<td>Sensitivity Analysis</td>
<td>26</td>
</tr>
<tr>
<td>4 IMPLICATIONS AND FUTURE RESEARCH</td>
<td>28</td>
</tr>
<tr>
<td>Management Strategy</td>
<td>28</td>
</tr>
<tr>
<td>Statistical Application</td>
<td>31</td>
</tr>
<tr>
<td>5 CONCLUSION</td>
<td>36</td>
</tr>
<tr>
<td>REFERENCES</td>
<td>38</td>
</tr>
<tr>
<td>BIOGRAPHICAL SKETCH</td>
<td>41</td>
</tr>
<tr>
<td>Figure</td>
<td>Description</td>
</tr>
<tr>
<td>--------</td>
<td>-----------------------------------------------------------------------------</td>
</tr>
<tr>
<td>2-1</td>
<td>Suspect properties in Alachua County, overlaid with 1 x1 mile grid.</td>
</tr>
<tr>
<td>2-2</td>
<td>Suspect properties in highest density grid cell, East Gainesville.</td>
</tr>
<tr>
<td>2-3</td>
<td>Dot density map of suspect properties.</td>
</tr>
<tr>
<td>2-4</td>
<td>Dot density of transactions under 90 days.</td>
</tr>
<tr>
<td>3-1</td>
<td>Density grid using $w=20%$.</td>
</tr>
<tr>
<td>3-2</td>
<td>Density grid using $w=40%$.</td>
</tr>
</tbody>
</table>
THE GEOSPATIAL FOOTPRINT OF MORTGAGE FRAUD: A CASE STUDY IN ALACHUA COUNTY

By

Austin Evans

August 2011

Chair: Grant I Thrall
Major: Geography

Mortgage fraud occurs when one or more parties knowingly make deliberate misstatements, misrepresentations, or omissions during the mortgage lending process. In 2005, Florida, Utah, and Georgia exhibited the highest rates of mortgage fraud. Because of the competitive nature of the mortgage industry, lenders are forced to create many new products which are often susceptible to fraud. Some analysts say that one major cause of the increase in fraud is poor underwriting practices. While much of this can be attributed to faulty appraisals, misrepresentation on the loan application is a frequent occurrence.

Over $10B in possible mortgage fraud was been detected by investigative journalists of the Sarasota Florida Herald-Tribune with the assistance of Grant Thrall. Within the Herald-Tribunes trade area of Sarasota and Manatee Counties alone a possible $500 million of possible mortgage fraud was detected. Geospatial analysis and geospatial technology (GIS) adds to risk management procedures of financial institutions. Such analysis is not new to risk reduction. Financial risk managers rely upon credit risk models, address-level computer assisted mass appraisal pricing systems, automated default tracking. Together this provides an early warning system
which if followed can reduce mortgage default attributable to credit risk. Implementation of credit risk modeling is routine among financial institutions. However, the great amount of prospective mortgage fraud transactions revealed in the Florida study indicates that credit risk modeling is not effective in detecting the possible occurrence of mortgage fraud. We believe that it is not a lack of lenders’ desire to mitigate the risk of fraud; rather, the procedural methodology for an early warning system to reveal possible mortgage fraud has not been adopted.

We present a geospatial procedure that can be automated to set flags on transactions that are possibly fraudulent. Our procedure is supported by 65,000 sales transactions for the County of Alachua (Gainesville), Florida.
CHAPTER 1  
MORTGAGE FRAUD AND GEOGRAPHY  

Summary  

The *Herald-Tribune*, in a 2009 study, uncovered $10 billion in suspected mortgage fraud activity in the State of Florida. Analysts at the *Herald-Tribune* studied residential sales transaction data from 2000 to 2008. It is alleged that Florida was a leader in mortgage fraud instances during this time period.  

Mortgage fraud is an important issue right now because in 2009, it cost the mortgage industry an estimated $14 billion in losses. These losses will be passed on to borrowers through higher interest rates and lower property values. An example of fraud perpetration is when a property has been purchased and then sold within a short period of time with an unnatural price increase. It has been thought that collusion usually occurs between the buyer, seller, and appraiser. In the appraisal practice, anything short of an arms-length appraisal is regarded as an unacceptable appraisal. Under the Uniform Standards of Professional Appraisal Practice (USAP) Ethics Rule, an appraiser must perform assignments with impartiality, objectivity, and independence, and without accommodation of personal interest (USPAP 2010-2011). In such fraudulent circumstances, the loan may be defaulted on and the agents involved in the transaction keep the proceeds.  

The *Herald-Tribune* analysts were given advice from University of Florida Professor Grant Thrall, who specializes in Geospatial Analysis of Real Estate and Land Economics. Following his advice, analysts combed through the statewide database to identify possible fraudulent transactions.
The first pass of identifying potential fraudulent transactions was to select from the large database, those housing transactions that sold two or more times within a three month period, and increased in sales price by 30% or more. 30% was judged by analysts and Thrall as an appropriate cutoff because it was viewed as highly unusual to increase by more than that amount within such a short time period. It is likely that a lower percentage would have revealed a greater amount of mortgage transactions that represent possible fraud, but also a greater number of non fraud transactions. In the context of this demonstration analysis, the analysts decided to use a more conservative figure. The statewide database was initially reviewed, but since the Herald-Tribune’s trade area is Sarasota and Manatee Counties, the decision was made to focus on these counties for local relevance. However, any significant occurrences of suspected fraud that were revealed were sent to federal regulators for review.

The Herald-Tribune study was published on July 19, 2009. In their study of the two counties, about 2,500 real estate transactions representing more than $500 million were highlighted as fraud. The unpublished findings of the entire State of Florida show that this is a statewide phenomenon. The question I am addressing is: Is there evidence that Alachua County has been impacted by similar fraudulent transactions, or is Alachua County unique in not having similar indicators? This is not only a statistical process, but utilizes GIS databases for management procedures. The geospatial analysis came to the decisions as to where additional attention should be given. If a transaction is highlighted, it is advised that further review should be conducted. Currently, there is nothing in place to flag suspicious transactions. There are multiple flags used when looking at individual loan documents, but nothing to filter transactions.
I am recommending a procedure similar to the *Herald-Tribune* study which could be refined to create better ratios of inclusion and exclusion, meaning fewer non-fraud transactions included and fewer fraud transactions excluded.

The *Herald-Tribune* study had four steps which will be replicated in my analysis. The first step is to obtain the entire property sales transaction database for Alachua County, which includes all transactions from 2000 through 2008. This data is available from the Alachua County Property Assessor’s website. The second step is to include only residential transactions that conform to the criteria of 30% or greater change in market transaction value within a three month time period. Of the over 65,000 residential property transactions in Alachua County over this time period, 424 fit our criteria. The third step is to identify commonalities between suspected mortgage fraud transactions. Commonality can be the same actors involved in multiple transactions, or concentrations of transactions in small areas. The fourth step involves investigating the public records of individuals involved, or identifying if the transactions are spatially dispersed. The idea is that perpetrators either mine certain subdivisions or condo units in executing fraud, or they pick spatially dispersed properties to avoid attention.

The ultimate goal of this study is to outline a geospatial procedure that can be used by regulators to identify possible mortgage fraud occurrences. The vision is to create a system similar to that in which the IRS flags tax returns as suspicious. This procedure will need to be further developed by experts, but it is an outline of a management strategy that will significantly increase the efficiency of mortgage fraud investigations.
Background

Over the past decade there has been a lot of turmoil in the real estate industry due to the housing boom and subsequent burst. It is widely known that the mortgage industry had a lot to do with these trends. Subprime mortgage lending has been blamed for most of the problems in housing and accessibility to home loans. However, another major issue that has not received as much attention is mortgage fraud. Mortgage fraud can be defined as a crime whereby property values are artificially inflated in order to fraudulently remove funds from the property’s equity. When mortgage fraud occurs, one or more parties knowingly make deliberate misstatements, misrepresentations, or omissions during the mortgage lending process (Carswell & Bachtel, 2007). This has become a serious problem in Florida, yet there is often a lack of law enforcement attention given to these crimes. In smaller cities and municipalities, the majority of law enforcement funds and time are spent on more violent, street level crime. The time necessary to investigate white collar crimes is also a burden on local law enforcement (Friedrichs, 1996). The goal of using GIS technology is to make fighting these crimes easier for law enforcement from a time and funding standpoint. GIS can help with data collecting, tracking occurrences of fraud, and preventing future illegal transactions. The aim of this project is to utilize GIS in expanding on a previous study of mortgage fraud in Florida. The previous study, completed by the Herald-Tribune in Sarasota, FL, found an estimated $10 billion worth of fraudulent transactions across the state. The study was then focused on Sarasota and Manatee Counties, in which $500 million of fraud was discovered. The intent of my study is to expand on this two county study and see if the results can be replicated in Alachua County, and study the geographical trends of ‘suspect’ transactions.
Risk management, or the lack there of, was one of the main factors in the demise of the mortgage industry in recent years. Lenders were more focused on maximizing Return on Equity (ROE) with little attention to Risk Adjusted Return on Capital (RaRoC), which leads to investment decisions being made without regard to the riskiness of the assets. Short term incentive structures in place for managers created lowered loss aversion among lenders, encouraging management to follow competitors into more risky products (Rossi, 2010). Rossi concludes from his study a number of important lessons for the mortgage industry, regulators and investors:

- the need for a comprehensive focus on development of industry-wide data and techniques for measuring risk; implementation of risk-adjusted return measures for firm objective setting; individual performance assessment; and greater introspection on the part of management teams to validate external information against a stated risk vision (Rossi, 2010).

There has been much written about the use of GIS in the real estate, appraisal, and mortgage industries. However, the application of this technology to criminal activity like mortgage fraud is a relatively uncharted field. For years mortgage lenders have used GIS to evaluate markets. The segregation of residential space into discrete geographic submarkets influences the pattern and nature of mortgage product demand. Neighborhood conditions affect credit risk because of their influence on housing prices (Belsky, Can, & Megbolugbe, 1998).

The USDA is currently using Landsat satellite imagery as a fraud detection tool (Rocchio, 2006). Crop insurance fraud is committed by reporting crop damages or losses when no such loss was incurred, usually because the crop was never planted. The Landsat images are updated bi-monthly, and save the U.S. Government approximately $100 million annually (Rocchio, 2006). When a suspect fraudulent claim is made, analysts refer to the image dataset to determine if in fact the subject crop was
planted, and if so whether or not it was actually damaged. This process allows cases to be settled quickly, saving the government millions in litigation costs and payouts. The fraud mitigation strategy employed by the USDA is an example of how geospatial analysis can streamline the investigation efforts of a regulatory agency.

Risk management is one of the major benefits of GIS to the mortgage industry. In order to reduce the risks associated with lending, lenders use local credit risk modeling, automated address-level pricing systems, automated default tracking and mapping systems, automated early warning, and local default systems based on automated appraisals (Belsky, Can, & Megbolugbe, 1998). While fraud is not as big of a risk as credit risk, prepayment risk, or default risk, it seems that institutions should still be concerned with limiting their exposure to this type of risk. However, the problem is not the lenders’ desire to mitigate the risk of fraud; it is their lack of ability. This is due to the minimal research in the field of mortgage fraud.

Based on 2005 statistics, Florida, Utah, and Georgia are the states with the highest rates of mortgage fraud. Because of the competitive nature of the mortgage industry, lenders are forced to create many new products which are often susceptible to fraud (Sharick, Omba, Larson, & Croft, 2006). Some analysts say that one major cause of the increase in fraud is poor underwriting practices. While much of this can be attributed to faulty appraisals, misrepresentation on the loan application is a frequent occurrence (Pendley, Costello, & Kelsch, 2007).

The idea of using maps and geospatial concepts to reduce general mortgage risk is not an entirely new concept. The Home Owners’ Loan Corporation created color-coded maps for cities across the country between 1935 and 1940 that indicated risk
levels for long-term real estate investment (Hillier, 2005). While the Home Owner’s Loan Corporation was created to reduce residential foreclosures, the organization began a project to investigate real estate conditions in cities across the country by creating a series of maps for 239 different cities. The maps were created to graphically reflect the trend of desirability of neighborhoods from a residential viewpoint (Federal Home Loan Bank Board, 1937). The maps represented residential areas with four different colors that described the market conditions of the neighborhood. The worst neighborhoods with poor demographics and little room for real estate growth were labeled red, which was a signal to avoid lending in that area. Even though this was a very rough form of geospatial risk management, it was a starting point for geographers and economists to build on.

There are some recent steps being taken across the country to understand and mitigate mortgage fraud, although most lack a geospatial component and often fail to effectively minimize fraud occurrences.

For example, Indiana recently created a task force designed to monitor mortgage lending and aid in fraud prevention. This task force is made up of state employees across several agencies that are responsible for combining the departments’ reports, policies, and recommendations on mortgage fraud (Zitrin, 2009). Zitrin goes on to suggest that this is only a start, and that a more appropriate effort would be to conduct investigations and mortgage application audits similar to the manner in which the Internal Revenue Service audits federal tax returns. Zitrin also describes the federal government’s half hearted attempt to battle fraud by enacting PL 110-289, which requires the reporting of confirmed or suspected fraudulent mortgages on a case by
case basis. This fails to address fraud before or while it is occurring, only the investigation of known cases. It seems more appropriate to implement a strategy to fight this activity before it happens by monitoring mortgage applications. Currently, a lender can lend to a buyer with no attempt to discover if there is dishonesty in the application, and no duty to report such discovery if one is made (Zitrin, 2009). A more strictly policed mortgage application process would inherently lead to fewer attempts to commit fraud.

In February 2011, New York City launched a computer program designed to flag suspicious properties in the city’s public property database (Buckley, 2011). This program, implemented by the mayor’s Financial Crimes Task Force, searches the cities property database for indicators of fraud like multiple changes in title, the transfer of titles at below-market prices; and the sale of properties at prices beneath the minimum amount required for tax filings. The properties that are flagged will then be sent to the District Attorneys’ Offices, the Police Department, and the Department of Investigation (Buckley, 2011). This new monitoring system has come in the wake of record amounts of mortgage fraud in New York. New York was ranked second in the nation behind Florida in total mortgage fraud cases in an April 2010 report (James & Butts, 2010). This is an example of how important mortgage fraud is to the real estate industry. Government regulators are beginning to see the importance of being proactive in fighting fraud as opposed to reactive. The only way to prevent fraud and protect victims from losing thousands, or perhaps millions, is to actively monitor transactions and prevent these fraudulent occurrences.
The Financial Crimes Enforcement Network has recognized the growing problem of mortgage fraud, citing in a 2006 report that Suspicious Activity Reports (SAR) increased by 1,411 percent from 1997 to 2005 (Mortgage Loan Fraud: An Industry Assessment based upon Suspicious Activity Report Analysis, 2006). According to the FinCEN report, fraud is growing because “it can be very lucrative and relatively easy to perpetrate, particularly in geographic areas experiencing rapid appreciation.” While these SAR reports are speculative and not confirmed cases of fraud, it is a substantial growth in suspicious activity. Everyone from appraisers, lenders, and brokers have been so caught up in the growing market and increased fees, they failed to seriously investigate these mortgage applications. If a buyer was willing to pay an abnormally high price, it meant more commission or fees for the participants in the transaction and so they had no incentive to question whether these transactions were arms length.
CHAPTER 2
ANALYSIS

Data Filtering

The Herald Tribune focused their study on Sarasota and Manatee Counties, and uncovered more than $500 million of highly suspicious transactions. While the number is likely lower for Alachua County, it will still be significant and will be a good sample area for a geospatial extension of this fraud detection process. The first step of this process is to replicate the statistical analysis that the newspaper used.

The initial data set consisted of over 19 million records pulled from various property appraiser data sets from across the State of Florida. Data was collected in the summer of 2008 from 57 of Florida’s 67 counties, and included property transactions dating back to 2000. Below are the relevant data fields:

- Parcel ID Number – essential for joining the dataset to Alachua county parcel layer and making in geospatial
- County – county in which the property lies
- Seller/Buyer – seller and buyer were provided for most counties, including Alachua
- Book/Page – Book and Page in which the deeds were recorded
- Amount – Purchase price of the property
- Date – Date of the transaction

The initial data set was reduced from 8 million transactions to the 65,000 plus that occurred in Alachua County in the time period. From here, excel was used to sort and filter the data to arrive at what we believed to be ‘suspicious transactions’.

The transactions were first filtered to show only the properties that sold more than once during this time period. Next these properties were filtered again to leave only
those that sold twice or more in a three month period. The final step was to sort through these properties and find the ones that increased in value by 30% or more. Experts believe that even during the real estate boom, an increase in value of 30% within a three month window was more than double that of average price increases. In the 30 years from 1975 through 2005, the average national level of home value appreciation was 6% per year. In high growth periods like 2005, that number was around 12.5%. In the coastal areas that saw the highest levels of appreciation, a 20% annual value increase was observed (Mortgage Loan Fraud: An Industry Assessment based upon Suspicious Activity Report Analysis, 2006). These figures were all annual, which is why experts say that a 30% increase over a 90 day period is a conservative figure in detecting suspicious transactions.

Of the results, properties that transacted for $100 were thrown out as these were most likely some type of warranty deed transaction. Transactions that occurred on the same day were also thrown out due to the way some municipalities record deeds. The first sale could have occurred before the house or condo was built, and the second many months later upon completion.

What remained were 424 property transactions across Alachua County that all showed similar characteristics to fraud occurrences. This statistical filtering method is a significant step, and allows fraud investigators to focus their investigations on the most likely occurrences. However, we believe that adding a geospatial element to this analysis will enable investigators to concentrate investigations even more.

It is important to remember that none of the transactions in the study that we label suspicious are confirmed fraud. These are all ‘likely fraud’ based on several
characteristics that are consistent with mortgage fraud. These transactions must all be reviewed by law enforcement and government regulators to determine whether or not fraud has actually been committed. This study is simply designed to provide these regulators with an effective GIS management strategy to reduce the time and effort that is spent on fraud investigations.

**Making the Data Geospatial**

The transaction data sets obtained from the Alachua County Property Appraiser contained the Parcel Identification Number, which allowed us to easily geocode the properties. A parcel layer of the entire county is available for download from the Property Appraiser’s website. Once this layer is added, the excel database of properties can be added and then joined to the parcel layer using the Parcel Identification Number (PIN) to match the data with Alachua County parcels.

The initial join required to match the parcel layer and the data is a one to one join. This means that each row of data, or property, is joined with one particular parcel. The one to one join provides us with the location of all property transactions that are labeled ‘suspect’.

**Grid Cell Approach**

The first method employed to detect trends and patterns in the fraud activity was to create a grid cell overlay. The idea is to create a one mile by one mile grid to overlay with the properties, and investigate the grid cells with the highest density of properties. Of the top three densest cells, two were located in central Gainesville and one was located in Alachua just south of Hwy 441. These three grid cells contained 10 or more properties in them, most of which were in the same neighborhood in all three cases.
This method provided us with 36 suspicious deals that were grouped together in similar neighborhoods.

It seems obvious when studying spatial data that you should also investigate the surrounding observations as well. In order to determine if less than 90 days is a reasonable timeframe to suspect fraud, we looked at the population of the entire grid cell. When looking at the average number of days between sales, each of the three cells had a population average over 250. It seems reasonable to assume that if the average time between sales for properties in a neighborhood is over 250 days, properties that transact multiple times within 90 days are out of the ordinary. This helps assure us that the time constraints used in our statistical analysis were reasonable.

Figure 2-1. Suspect properties in Alachua County, overlaid with 1 x1 mile grid

When looking at the grid cell in Alachua above, it is easy to see the correlation of suspected fraud and transactions with less than a 90 day time period. In this case there
are three properties that sold within the 90 day time frame, yet did not get flagged as possible fraud because they didn’t also match the price increase constraint. This tells us a couple things. First, the fact that we used multiple constraints in flagging suspect properties was essential. Here are three properties that would have been included in our final data set had we used only a time constraint. It also provides a snapshot of the other sales in the neighborhood. We can see that the other properties sold over a significantly longer time period, with eleven of those being over 225 days.

Figure 2-2. Suspect properties in highest density grid cell, East Gainesville

**Dot Density Approach**

The grid approach is effective in highlighting the grid cells with the highest concentration of observations. The fault in this method lies in the fact that these cells have boundaries. In some of the densest cells, the borders cut through a neighborhood or subdivision. Even when altering the location of the grid, there is no way to
encompass an entire hot spot without breaking it up into multiple cells. To avoid this, the dot density method can be employed. By locating each observation as a point rather than a polygon, a density map can be created to show high density hot spots without boundaries.

Figure 2-3. Dot density map of suspect properties
Figure 2-4. Dot density of transactions under 90 days

The idea is that with a dot density map we can focus on an entire hotspot. In figure 3 above, you can see that there are two major concentrations of suspect properties. One occurs in Central Gainesville and one occurs in the City of Alachua. The grid cell approach basically breaks these areas up into grid cells of a certain density. The dot density approach displays the center of the hot spot and the direction and concentration in which it expands from that center. This is more valuable to an investigator analyzing an area or neighborhood.

When looking again at Figure 2-3, it is easy to locate the two hotspots and the neighborhoods that they occur in. An investigator can easily select these neighborhoods to get property information in which to investigate. This is a more streamlined process than looking over several grid cells and selecting properties from the three or four different grids.
As previously discussed, we identified 424 suspected fraud transactions in Alachua County between 2003 and 2008. These transactions represented $72,000,000 in total value (TV) of suspected fraud for the five year span, which equates to an average sales price of $170,000. It is hard to determine an average down payment percentage for these transactions, but based on the Florida housing market during this time period 5% is a reasonable figure. Traditional down payments are 20%, but in recent years government backed loans and creative financing options have made a 5% down payment more commonplace (Down Payments and Closing Costs, 2011). Subprime mortgages allowed borrowers to pay as little as 3% down from both private and public institutions.

Based on a 5% average down payment on $72,000,000 worth of transactions, we can calculate a potential fraud amount of $68,400,000.

Potential Fraud = (1-DP%) x TV

If we use as average down payment of 10%, we calculate a potential fraud value of $64,800,000. Both of these numbers are significant and a cause for concern.

**Gaussian Smoothing**

The density grids are a way to visualize point data by transforming individual points into a grid. Each of these grid cells are then given a value determined by the density of nearby points. In the Alachua parcel maps I used the Quartic method. This approach uses a quartic approximation of a Gaussian distribution to apply factors to the weighting values. In each map, I created a grid based on the selection set in the layer. The first selection is the centroid points from all of the parcels that transacted twice.
within a ninety day period. The second selection is the points from all of the ‘suspect’ parcels.

Maptitude uses a Gaussian Smoothing technique in its dot density function. The Gaussian smoothing operator is a 2-D convolution operator that is used to blur images and remove detail and noise. It uses a kernel that represents a bell shaped hump. The Gaussian creates a weighted average of each of the subject pixels’ neighborhoods, where the average is more highly weighted toward the central pixels. This provides better smoothing than a traditional mean filter which uses an even weighted average.

**Sensitivity Analysis**

The variable $w$ used in the study was set at 30% based on input from experts and analysts. This figure was deemed to be significantly greater than sales price increases at the time. However, a sensitivity analysis was performed in order to examine how a larger or smaller percentage price increase could affect the number and distribution of suspected fraud transactions. $W$’s of 20% and 40% were used to filter the transaction database in order to perform the sensitivity analysis.

As expected, the 20% and 40% $w$’s resulted in increased and decreased data sets, respectively. The $w=20\%$ analysis resulted in a suspected fraud data set of 385 transactions. The $w=40\%$ analysis resulted in a suspected fraud data set of 496 transactions. This is logical that tightening the constraints of the filter result in decreased observations, and loosening the constraints result in increased observations. However, the spatial distribution of the hot spots remained relatively constant. The main hot spot in East Gainesville is present in all three variations of the $w$ variable. The neighborhoods with high concentrations of suspected transactions remain the same in all three analyses; it is the scattered observations that differ in each case.
Figure 3-1. Density grid using w=20%

Figure 3-2. Density grid using w=40%
CHAPTER 4
IMPLICATIONS AND FUTURE RESEARCH

Management Strategy

The purpose of this application of GIS analysis is to create a strategy that can be used to improve business and management strategies. The end user of this process would ideally be an analyst with a government regulator or an analyst with a mortgage lender. The four step process would be implemented in this way: (1) obtain entire property transaction database from Alachua County Property Assessor, (2) filter database to include only residential transactions that increased in market sales price of w% or more within 90 days, (3) identify commonalities in actors involved or in spatial pattern of properties, and (4) reveal whether public records or spatial data shows patterns in actors or location in the suspicious transactions.

The first step in this management strategy involves obtaining the entire residential property transaction database for the desired time period. The time period used in this study was 2003-2004, which captured the height of the real estate boom in Florida and subsequently the height of the suspected fraud instances. Transaction databases are available from most property assessors’ websites, which allow users to sort transactions based on property type, date of transaction, seller/buyer, and several other categories. This is an appropriate strategy when gathering data on a weekly or monthly basis. However, when collecting initial data sets or data sets over a longer period of time, it is recommended that users obtain this directly from the property assessor’s office. The aim would be to build an initial database and maintain it on a regular basis.

The next step of the process is the filtering of the property transaction database. In this study the data was input into Excel, and then sorted and filtered to create a list of
suspected fraud transactions. We were analyzing a historic data set, where as ultimate users of this process would be dealing with a more fluid data set. Once a database of the desired market is created this would be constantly updated with current transactions. In this case, a geospatial database would be more efficient than a spreadsheet like Excel. This goes beyond the scope of this study, but ultimately a continually changing geospatial database would sort new transactions and flag them as suspect or not as soon as they are input. Transactions that are flagged would automatically be mapped on the current dot density map. This would keep the hot spots up to date and current.

Once a geospatial database is in place, analysts will be able to identify commonalities and patterns in both actors in the transactions and spatial distribution. Identifying links in actors involved would be done by searching the updated transaction database as well as public records. Thing such as same name, company name, and mailing address that appear on the deed would be a red flag. Often times, investors will use an LLC or partnership to buy the subject property. This is to try and limit their exposure, but patterns and links in ownership can be found. Also, investors will use different company names on the recorded deed to try and defer attention. This can be addressed by searching the mailing address, as similarities in PO Box or address might point out connections in investors.

The main focus of this study is the geospatial analysis of commonalities. In order to understand how instances of possible fraud are patterned, we must map them and interpret whether they are spatially conglomerated or dispersed. Going back to the
geospatial database mentioned above, new occurrences of possible fraud will be input routinely into this database to achieve two desired results.

For one, this will ensure up to date maps and dot density grids in order to have ever changing and highly accurate hot spot analysis. It is recommended that this be done on a weekly or monthly basis. Analysts would then be able to search a desired real estate market and determine where the present potential fraud hot spots are occurring, and how they have shifted over time. It is likely that while the five year period analyzed in this study pointed to distinct hotspots, if the suspected fraud occurrences over the past 20 years were mapped on a yearly basis you would see a definite migration.

In the case of Alachua County, hot spots occurred in less affluent neighborhoods. This could be either a cause or an effect. It is widely known that mortgage fraud negatively affects neighborhoods by decreasing home values. This process takes several years though, and the neighborhoods discovered in this study have been in their present state since well before the suspected fraud took place. The better explanation is that perpetrators seek out lower end neighborhoods due to ease of access. It is easier to get a mortgage for a smaller amount than a large amount, and so investors would be attracted to areas with lower property values.

The change in markets and submarkets over time would imply a similar shift in suspected fraud hot spots over time, only lagged. For instance, the housing market in Gainesville has steadily moved west over the years. The higher end communities have shifted west from central Gainesville to the western edge, along the I-75 corridor. This means the lower end communities and neighborhoods are now in central and east
Gainesville, where one of our hotspots can be found. While these changes are slow moving and don’t have much effect on day to day suspected fraud occurrences, they can be instrumental in modeling and projecting future hot spots.

The second major benefit of maintaining an up to date geospatial database is that transactions which aren’t initially flagged as suspected fraud will be flagged if they are located in known hot spots. For instance, a perpetrator that is participating in appraisal fraud as opposed to flipping fraud could be flagged in this case. This transaction could be the first sale of this property in several years, and so the significant price increase wouldn’t initially be caught due to the long period between sales. However, the investor colluded with an appraiser and received an above market value. This appraisal was then taken to a bank where an inflated mortgage was obtained. If this property was mapped and located in a hot spot, the transaction would be flagged as suspected fraud and reviewed further. Otherwise it would not be caught until the mortgage was defaulted on and the lender was left with a property worth less than they loaned to the investor.

**Statistical Application**

Beyond just the geospatial characteristics of transactions, statistical analysis at the neighborhood level would positively enhance the model. A test such as Grubb’s Test for outliers would be an additional filter in creating a more desired inclusion verse exclusion ratio. The geospatial component of this model patterns suspected fraud occurrences into hot spots, and points out neighborhoods and submarkets that are conducive to fraud. However, an application such as the Grubb’s test could predict outliers at the neighborhood level. The idea of property value trends differing across neighborhoods and submarkets could be accounted for with this type of test.
Grubb’s test for outliers detects observations that are significantly larger than the mean in a normally distributed data set. The test is designed to identify one outlier at a time. Once an outlier is detected, the observation is removed and the test is repeated. A standard Grubb’s Test is shown in Equation 4-1:

\[ G = \frac{Y_{\text{Max}} - \bar{Y}}{S} \]

In this equation, \( Y_{\text{Max}} \) represents the largest observation, \( \bar{Y} \) represents the population average, and \( S \) is the standard deviation (NIST/SEMATECH e-Handbook of Statistical Methods). This equation could be modified to detect outliers in a neighborhood or submarket. The one sided test would be preferred in this case because we are searching for excessive appraisal values, and so only maximum outliers would be desired.

\[ G = \left| \left( \frac{APV}{ASV} \right) - \left( \frac{APV_{\alpha}}{ASV_{\alpha}} \right) \right| / S \]

In this case, APV represents appraised value, ASV represents assessed value, and \( S \) is the standard deviation of the ratio (APV/ASV). The \( \alpha \) in this equation refers to the median APV and median ASV. These median values would be calculated for neighborhoods or submarkets, as opposed to city or county wide values. Since some areas and submarkets experience higher growth than the city average, basing this ratio on area specific values would be more appropriate. The Grubb test being implemented here is designed to detect outliers by comparing assessed value to appraised value.
ratios. The calculated G statistic\textsuperscript{1} would be compared with a critical value based on the predetermined alpha level. Observations with G statistic values greater than the critical value would be considered outliers.

Appraised value (APV) is an appraiser’s opinion of the property value, which is utilized to obtain a mortgage. The appraised value is what a professional believes the property is worth in the current market. This value can be slightly volatile, and is also the value in which fraud is based. Inflated appraised values are based either on the straw buyer’s price that they are willing to pay, or appraisers conspiring with the fraud perpetrators. The assessed value (ASV) is the county property assessor’s opinion of value, from which the property taxes are calculated. This value is more conservative and usually is more consistent across the market. Assessed values are calculated once a year, whereas appraised values can be calculated multiple times per year. When a mortgage application is being processed, banks require a certified appraisal to ensure the mortgage request is consistent with the value of the property.

The ratios of APV to ASV in a submarket or neighborhood should all be similar. The ASV’s will be pretty standard since the local property assessor’s office utilizes the same procedure to value all residential properties. The appraisal practice is a consistent process across the board due to strict Uniform Standards of Professional Appraisal Practice (USPAP) guidelines. This means all professional appraisers are held to the same standards when deriving an opinion of market value. Significant outliers in APVs are likely the result of appraiser error or unethical practice.

\textsuperscript{1} The G statistic, G, in this paper refers to Grubb’s statistic. This is not to be confused with the Getis-Ord G statistic.
Due to the expected consistency in both APV and ASV, it is reasonable to assume the ratios to be consistent across submarkets. For example, if a submarket has an average APV of $140,000 and an average ASV of $100,000, the resulting ratio would be 1.4. When looking at individual properties in that submarket, a ratio within a determined standard error would be expected. When a property with an unusually high ratio is plugged into the Grubb equation, the resulting G statistic would be above the critical value and be recognized as an outlier.

This type of Grubb’s test analysis would be an ideal complement to the geospatial element of this model. As mortgage applications are processed, they would be run through the model and immediately flagged if the G statistic was significant. This would allow further underwriting to be performed before the mortgage was granted, helping to prevent fraudulent transactions. Obviously all properties that appear as outliers in the Grubb’s test are not fraud, but are consistent with characteristics of fraud.

Acquiring assessed values for this analysis can be done by simply visiting the county Property Appraiser’s website. Current year assessed values can be obtained for any property in the county. Appraised values would be included in mortgage applications, so whether this analysis is used by lenders or regulators access to the appraisals wouldn’t be an issue. Ideally, databases will be maintained for individual submarkets. Every time a mortgage application gets submitted, the G stat will be calculated and then stored in the database. If it is an outlier, it will then be flagged and further investigated. If not, the ratio will simply be added to the submarket database to update the median ASV and APV.
As previously mentioned, appraisal fraud is one of the major types of mortgage fraud being committed. Investors seek an above market appraisal so that they can acquire an inflated mortgage. In such case, the inflated appraisal would create a high G statistic and flag the property as an outlier.
Mortgage fraud has become an increasingly popular white collar crime due to its high returns and fairly week regulatory practices. The damage from these crimes is felt by more than just the lender who is left with a defaulted loan. Neighbors and others in the market or submarket in which this activity is occurring are victims of decreasing home and property values. These areas see high rates of foreclosures and short sales which lead to a high vacancy rate and decreased property values. This can then be compounded when these neighbors recognize a below market home value and find out that they are underwater on their mortgage.

Appraisers are also affected by fraud schemes. When selecting comparables to use in an appraisal, fraudulent transactions can show up and lead to a faulty appraisal. Without regulation, innocent appraisers are unsure which comparables are market value and which are artificially inflated.

It is essential that regulators step up their mortgage fraud detection and prevention efforts. It is understandably difficult to determine and prevent fraud, but the implications of not making it harder to commit are currently being seen. What we have addressed in this study is in no way the answer, but merely an outline of what investigation procedures should look like.

Like real estate, mortgage fraud differs from market to market. The ideal strategy for monitoring fraud seemingly would involve analysis at the submarket or market level, whether that is by county, MSA, or region. The process outlined in this paper would best be utilized in such scenario, where geospatial databases are maintained for each submarket. The analysts in each market would then have an enhanced knowledge of
local trends and patterns, much like a broker or appraiser focuses on a particular geographic market.

This strategy would allow analysts to monitor trends and shifts in suspected fraud activity, and better prevent future occurrences. A GIS would need to be developed to incorporate a spatial database that filtered and mapped suspect transactions. This GIS would also flag suspicious transactions that occur in known hotspots.

GIS has been instrumental in real estate analysis in the private and public sectors, but few have fully utilized its white collar crime prevention capabilities. In this study we were able to take 65,000 property transactions, filter them to 424 suspected fraud transactions, and then locate hot spots and high risk subdivisions in Alachua County. A process similar to this would significantly increase the efficiency of mortgage fraud investigations. This is an initial step, but experts should be able to fine tune the process based on their knowledge of the industry and the crime. Whether it’s adjusting the price increase parameters, time frame, or density that is considered significant, a finely tuned system such as this would be leaps and bounds above what is currently in place.
REFERENCES


BIOGRAPHICAL SKETCH

Austin Thomas Evans was born in 1986 in Lakeland, FL. The oldest of three children, Austin grew up in Lakeland and graduated from Lakeland High School in 2004. He earned his B.S. in finance from the University of Florida in 2008.

Upon graduation from the University of Florida in 2008, with his B.S. in finance, Austin began working at LM Properties in Lakeland, FL. While working in land development and finance, he developed a passion for real estate analysis and wanted to further his knowledge. Having taken undergraduate classes with Dr. Thrall in GIS for real estate market analysis, Austin decided to pursue a master’s degree in geography with a minor in real estate from the Hough Graduate School of Business.

Upon the completion of his M.A. degree from the University of Florida, Austin began work at Land Advisors Organization in Orlando, FL as an Analyst and GIS Specialist.