THE DESIGN AND DEVELOPMENT OF WEBCAT: AN ONLINE CRIME ANALYSIS TOOL

Student team: Calvin Francis, LaLisha Hurt, Joseph McCormick, Nicola Palmer

Faculty Advisor: Donald E. Brown, Department of Systems Engineering

VIJIS Staff Advisor: Jason Dalton

Client Advisors: Butch Johnstone and Chuck Ruble
Virginia Department of Criminal Justice Services
Technical Specialist Division
Richlands, VA
E-mail chuckr@netscape.net

KEYWORDS: Crime analysis, GIS, crime mapping, ArcIMS, SVG, NIBRS, XML

ABSTRACT

This Capstone team designed and developed an internet-based crime analysis system that provides modern spatial and temporal analyses and reporting over the web. Once implemented, this system will provide crime analysts with a powerful tool to more effectively use crime data to better allocate their resources to improve crime prevention and criminal apprehension. The team has met its objectives by successfully developing all the components of this system.

INTRODUCTION

Law enforcement agencies struggle with large amounts of data that needs to be turned into useful information. With the incredible advancements in computer processing power as well as memory capacity, it has become possible to analyze this data in a short period of time. The Virginia Department of Criminal Justice Services (VDCJS) wants to use this capability to improve the efficiency and effectiveness of law enforcement throughout the Commonwealth of Virginia. However, due to the cost, time, and resources needed for the implementation and upkeep of current systems, as well as the geographical location of certain jurisdictions, this has not become a reality.

Here at the University of Virginia in conjunction with the Virginia Institute for Justice Information Systems, our Capstone Team has worked to address the need of regional law enforcement as defined by VDCJS by creating the Web Crime Analysis Tool (WebCAT). In defining the requirements for regional crime analysis, we developed the following objectives:

1. To create a system accessible over the internet that allows for widespread data access
2. To design an easy-to-use interface
3. To analyze geographical trends in data over time
4. To provide statistical analysis
5. To generate reports for a jurisdiction’s internal and external use.

WebCAT is designed to overcome the limitations of current systems and provide agencies with a systematic method to better allocate their police resources. Implementing WebCAT in the Mountain Empire, a region in Southwestern Virginia, will help illustrate the importance of crime analysis and data sharing amongst different jurisdictions.

RELATED RESEARCH

National Incident Based Reporting System

The Commonwealth of Virginia is divided into 261 police jurisdictions. Each jurisdiction collects an abundance of crime data and inserts it into their database daily. The problem is that police officers only have direct access to the data collected by their jurisdiction. The only statewide sharing of this data is through the National Incident Based Reporting System (NIBRS), in which the incident data collected is filtered to the NIBRS format, saved to a collection of floppy discs, and mailed to the state NIBRS center monthly. Once there, the incident data is published in an annual statistical report for use by law and policymakers. Therefore the crime data analysis done by a law enforcement agent is limited to data on incidents that have occurred in his jurisdiction. Any cross-jurisdictional patterns are lost and may not be detected by an officer in either jurisdiction [VGIN].
ReCAP and CARV

Two applications have been developed by the Virginia Institute for Justice Information Systems (VIJIS): the Regional Crime Analysis Program (ReCAP) and Crime Analysis and Reporting for Virginia (CARV). These programs are capable of collecting statewide data and allowing statewide access to this data. They also analyze the data, plot the analyses, report the analyses, and map the data using GIS (graphical information systems) software. [Curry]

Despite their successful development, these two systems have limitations. ReCAP software is free to jurisdictions within Virginia, but it costs money to integrate their system with the ReCAP Database Management System. It also costs the jurisdiction money to maintain and update the system. The creators of CARV made the tool more affordable through a cost-effective and flexible design, but it only has 25 ports, which limits the number of persons who can use the system and inhibits the speed. [VGIN]

The Need for Web Capabilities

With the increasing capabilities of dynamic websites, other applications are quickly losing ground to internet tools which provide less hassle and cost for the client. Web-based applications maintain a centralized location that can be accessed by multiple users.

DESIGNING WEBCAT

System Metrics

Using the objectives shown in the introduction, the Capstone Team formulated the following metrics, the first being the most important:

1. Cost
2. Accuracy
3. Speed
4. Usability for client
5. Ease of Implementation

Website Interface and Architecture

We designed eight possible layouts of the website and chose the one with the most user-friendly interface and intuitive functionality. We created the site architecture to determine what components would be available throughout the website, how they would communicate, what data each would need, and the output they would generate.

Technology Alternatives

Database Originally the Capstone Team wanted to include a Database Management System (DBMS) as part of WebCAT. We decided, however, that a temporary database would increase the speed of the system by decreasing the amount of data transmission and storage. Instead of uploading all crime data collected, the user queries their own DBMS and uploads the resulting recordset into WebCAT. Jurisdictions in the Mountain Empire use the relational database Memex to store their data.

We chose the Extensible Markup Language (XML) to serve as the temporary database. XML is a markup language that uses tags to describe data and allow for its manipulation and retrieval. It stores data as text so that it can be shared across networks and used with any programming language. [Sall] XML is open source so it comes at no cost to the WebCAT team.

Web Development Tool After developing metrics for WebCAT, the Capstone Team selected alternatives for the Web Development Tool. These included Microsoft Visual Interdev, XML Spy 5.0, Cold Fusion, Front Page, and Macromedia Dreamweaver Ultradev. In order to select the most appropriate technology for each component of WebCAT, we ranked the alternatives based on specific metrics. We chose Microsoft Visual Interdev to implement the majority of the system because of its high score. However, for the report module, the team switched to Macromedia Dreamweaver Ultradev for use of additional features and ease of implementation.

Geographical Information System (GIS) The alternatives for the GIS were ESRI Arc IMS, ESRI MapObjects IMS, AutoDesk MapGuide, and Demis Map Server. ESRI ArcIMS was chosen because the University already owned the license so there was no cost involved. Also it has an easy to use interface and the latest GIS features, and was highly recommended by experts in the field.

Scalable Vector Graphics (SVG) We chose to use SVG to design the control and time charts for WebCAT. SVG is a type of Extensible Markup Language (XML) that allows for straightforward and efficient two-dimensional graphics on the web. With SVG, plain text is used to create graphics features that would usually be associated with bitmap images (jpg, gif). This greatly increases processing time. SVG also works well across all platforms, output resolutions,
color spaces, ranges of available bandwidths, and memory capacity. Finally, scripts can be embedded into SVG allowing for dynamic graphics. [Sall]

CRIME DATA IMPORT

A user can upload data into WebCAT following one of two methods. The first method involves allowing a user to import files such as Microsoft Excel (.csv files) and text tables. The second method allows the user to manually input (cut and paste) specific crime elements into a text area and submit that data.

Due to a user’s connectivity speed over the internet and server capacity, huge files cannot easily be uploaded. To achieve optimal performance, users are expected to only import those crime elements they wish to analyze with WebCAT.

WebCAT is designed to work with 20 predefined crime data elements. These elements were derived from desktop crime analysis applications and from correspondence with crime analysts. They are the 20 most universal crime elements. Each of these is an element in the XML document delineated by an opening and closing tag. The format used in this XML document is defined within an XML schema that we created. The purpose of a schema is to identify which elements and attributes can legally appear, what type they should be, whether an element can be empty, and what the default may be for that element or attribute.

By creating an XML Document Object Model within each WebCAT component we retrieved any element required for analysis.

STATISTICAL ANALYSIS

In order for jurisdictions to allocate their resources most effectively, they need to predict when they think crime will occur. To do this in WebCAT, they can use control and time charts. Control charts use the standard deviation of a set of crime incidents to detect when the number of occurrences is exceptionally high or low. This means that the control chart detects when crime is deviated above or below what is statistically normal. Time charts identify patterns that occur in crime during particular days of the week or times during the day.

Control Charts

In WebCAT, users will chose from two different control chart options. The user can choose to view their queried crime data in weekly bins or by individual days. Three lines are plotted on the chart. The centerline (CL) is the average or statistical mean of all the data. The upper control limit (UCL) is three standard deviations above the mean and the lower control limit (LCL) is three standard deviations below the mean. If a certain point is above the UCL or below the LCL, then that point is considered out of control, and the police should take note so they can take corrective action. Charts will always be drawn so that the UCL is visible.

Through the use of SVG, the Capstone Team was able to make the control charts dynamic so that the information changes depending on where the user moves the mouse. This ensures that no matter what the size of the intervals on the chart may be, the user will know the exact amount of crime committed.

Weekly Control Chart When used in previous crime analysis tools, the weekly control chart plotted the most recent 52 weeks of the year displaying how much crime occurred each week. The Capstone Team decided that it would be more beneficial to look at different years simultaneously so the user can see if trends occur at the same points each year. Each year is displayed in a different color. Figure 1 shows an example of a weekly control chart for robberies in Central Virginia during the years 1997 to 1999. Currently, the user has the mouse on week 36 of the year 1998.

![Weekly Control Chart](image)

**Figure 1: Weekly Control Chart**

90 Day Control Chart This control chart looks at the past 90 days beginning with the most recent date to the far right. This allows the user to see how crime has changed over the past three months by looking for points that were out of control. Since this chart is recorded on a daily basis, it gives a jurisdiction online process control. They can print the chart regularly and observe as crime changes over time. This chart is also dynamic so the mouse can be used to trigger information for each point.
THE Design And Development Of Webcat: An Online Crime Analysis Tool

Time Charts

The first time chart in WebCAT is on a radar plot. Figure 2 shows this chart with robbery incident data from 1997 to 1999 in Central Virginia. This time chart breaks time into 10-minute intervals over a 24-hour time period. This means that if data is queried for two months, all of the crimes committed within a range of 10 minutes will be clumped together so that the user can see whether to target one time of day versus another. The chart displays the mean and the standard deviation above the mean by using colored circles. If an interval during the day exceeds the standard deviation above the mean then that time is placed in the table to the right of the chart.

Figure 2: Radial Time Chart

The second type of time chart breaks crime down into days of the week. Every crime that is committed is put into the bin corresponding to the day of week that it falls on. The sum for each day of the week is found over the entire data set and a bar graph is formed. No form of mean or standard deviation is used in this chart. It simply allows the user to look at an aggregate bar graph.

SPATIAL ANALYSIS

Using ESRI ArcIMS, the Capstone team developed the WebCAT Dynamic Map Tool. This tool is served over the internet using HTML and Java 2 applets, which provide powerful GIS functionality and have the potential to provide a valuable crime analysis tool to law enforcement agencies once implemented into a crime analysis system. This dynamic map component presents the frequency of crime within a geographic region and will allow the crime analyst to determine the correlations between these incidents, such as date and time, or suspect description.

Map Service

The Dynamic Map Tool is fed data through the ArcIMS FeatureServer. The FeatureServer uses Java 2 applets to maintain a cache of the vectors, polygons, and objects viewed by the user on the client computer. Therefore, map manipulation is done in large part on the client side in order to improve response speed; as a request is sent to the server only when a new geometric feature is necessary.

GIS Functionality

The Dynamic Map Tool provides the user with a powerful GIS tool that performs spatial analysis on crime data. Upon initial load of this page, the user is given a view of Tazewell County, as well as the cities (labeled), rivers, and major highways.

The tool provides typical GIS functionality, such as zoom-in and zoom-out, to get a closer or broader view of the incidents plotted in the region. As the user zooms in for a closer look, more objects are revealed. At the scale of 1:100,000, the streets of Tazewell County appear; and at a level of 1:10,000, the names of the streets are labeled. These scales were chosen in order to avoid cluttering the map with too many items at one time.

Choosing the information tool, the user can click on an object of interest on the map and the Dynamic Map Tool responds with a table displaying the attributes of that object stored in the database. With this tool, the crime analyst will be able to choose an incident plotted and view its stored data.

The Dynamic Map Tool also provides several selection tools, rectangle, radius, line, and polygon, which allow the user to select more than one object on the active layer. Upon selecting several incidents using a selection tool, the crime analyst can view attributes of these incidents simultaneously by clicking the attributes tool. The analyst can then compare values, such as date and time, to see if perhaps a pattern emerges.

Lastly, the Dynamic Map Tool provides the user with a buffer tool and a querying tool, both with similar functionality. The buffer tool aids spatial analysis by querying the database for objects within an X mile radius of the selected incident(s), where X is set by the user. The map generates a new layer on the map, which overlays the X mile radius, as well as highlights the objects that reside within this radius (see Figure 3). The querying tool provides an interface through which users can generate a query, for example, if they want to view all incidents that were reported over the last 7 days.
Advanced Features

Perhaps the most significant GIS functionality offered by the Dynamic Map Tool is providing the use of ArcIMS MapTips. MapTips are small pop-up boxes that appear if a user floats the mouse pointer over an object on the map. In the developmental stage of the Dynamic Map Tool, the MapTips function was activated for all of the layers with line or point objects (i.e. the Rivers, Roads, and Point layer), setting them to display the name of the object. This is particularly helpful, for example, when the user zooms in to a scale where the Street layer becomes visible, but not close enough that the names of streets are displayed. At such a level, if the user wants to know the name of a street, he can simply float the mouse over that street and the MapNote will appear with the name.

The Java 2 applets featured in the Dynamic Map tool will allow each individual analyst to view different crime maps simultaneously. The applets provide an “Add Layer” function for the user. Selecting this function, the analyst can browse to the crime layer generated by WebCAT for him and add it to the map.

REPORTS

The reporting component of WebCAT addresses both the internal and external concerns of turning raw data into visual patterns. The purpose of this component is to provide users with an easy way to visualize crime data via segmentation. The system provides functionalities that allow agencies and police officials to group the data according to their needs.

After the user performs their analyses, officials can view their segmentation instantly. Since the data is already captured once the user starts the process, the module allows the user to filter that data. To segment out information not pertinent to the analysis, the user can search by crime type, criminal’s name, date range, or city. Once the query is selected, the recordset appears in a separate HTML page that can ultimately be printed, saved, or emailed for both internal and external use. Figure 4 shows an example of a generated report.

Figure 3: Dynamic Map Tool

Figure 4: Report Module

TEST PLAN

User Testing

We will perform numerous tests on our program. The main purpose of this testing is to identify any glitches and/or flaws of WebCAT immediately following implementation. One of the areas of concern will be identifying how well our system meets the user’s needs.

In conducting the user testing, the subjects will be police officers, students unfamiliar with crime analysis systems, Jason Dalton, and Dr. Donald Brown. Each subject will be given a list of tasks to perform and a survey to complete following their session. This testing will be used to gage how much feedback, visibility, and affordance our system provides, as well as test the system against chosen metrics.

Test Metrics

Before initial stages of design, we determined four metrics on which WebCAT will be evaluated:

1. Performance
2. Accuracy
3. Usability
4. Response Speed

Performance will measure the system as a whole, while accuracy will be an averaged score of each individual component. Usability will measure user-friendliness and affordance of the interface. Speed will measure the amount of time it takes WebCAT to respond to a user request. We feel that these metrics
will measure how well we have achieved our objectives for this project.

RESULTS

Though the WebCAT system is not fully implemented yet, we feel that what we have developed meets the objectives we set forth. All of the components have been created and are accessible on the World Wide Web. They all ensure accuracy and speed, and were designed at minimal cost. Following standard human-machine interaction principles, we were able to design an intuitive interface allowing for good usability.

Data can be placed into an XML document by either uploading a file or manually importing crime elements into a text area. The XML document follows the schema we created, which ensures that the data being imported fits one of the 20 predefined elements.

The geographical information system (GIS) developed combines maps and database management systems to generate a spatial analysis tool. This Dynamic Map Tool allows an analyst to query the map for information on reported incidents in analyzing geographical trends in the data.

The control and time charts were tested and modified using a variety of years and number of records. We tested the accuracy of the UCL, CL, and LCL, Mean, and STD calculations and their placement, the location of the points on the graphs, and the accuracy of the tables for all charts. We modified the code to generate these charts in the least amount of time possible since much iteration is required.

The report generator was cross-referenced with a temporary database to ensure that the search query resulted in precise data extraction.

CONCLUSIONS

With the combination of spatial analysis, temporal analysis, and reporting, the web crime analysis system developed by the Capstone team will provide analysts with a better understanding of the crimes reported over time within a geographical region. This system should aid law enforcement agencies to more effectively analyze crime and more efficiently allocate their resources in order to improve crime prevention and criminal apprehension. Hopefully, the use of WebCAT in the Mountain Empire will demonstrate its value and promote its deployment throughout the Commonwealth.

REFERENCES


VGIN Capstone Team. “Crime Analysis and Reporting for Virginia (CARV).” University of Virginia. Charlottesville, VA: April 1, 2001: 1, 7-8

BIOGRAPHIES

Calvin Francis is a fourth year Systems and Information Engineering student. He is currently employed by the government and will graduating next fall. His primary contribution to WebCAT was the development of the data importation system.

LaLisha Hurt is a fourth year Systems and Information Engineering from Richmond, VA. Her contribution to WebCAT was a report module. LaLisha will be working for the Federal Reserve Bank of Richmond as an information technology examiner upon graduation.

Nicola Palmer is a fourth year Systems and Information Engineering student from Moorestown, NJ. Her primary contribution to WebCAT was the creation of the time and control charts. Nicola hopes to be working in a boarding school in Europe next year.

Joe McCormick is fourth year Systems and Information Engineering student from Herndon, Virginia. His primary contribution to the project was the development of the GIS spatial analysis component. He will be joining Booz Allen Hamilton as a consultant later this summer.