WebCAT: Enhancement through Analysis of Additional Data

Eric J. Ast, Joshua E.W. Mines, Amrita Mukhopadhyay,

Donald E. Brown, James H. Conklin

Abstract -- Since the year 2000, the University of Virginia Systems and Information Engineering Department has been working with the Virginia Department of Criminal Justice Services (VDCJS) on WebCAT, the Web-Based Crime Analysis Toolkit. WebCAT provides users with crime analysis tools ranging from data query options to graphs and reports generation. Previously, WebCAT allowed users to analyze incident occurrences, such as robberies or kidnappings. The authors have expanded the analysis tools of WebCAT to include arrests, calls for service, offenses, people, and warrants data types. The addition of these data types provides users with analysis capabilities beyond the previous version of WebCAT. For each additional data type, the authors adapted previous WebCAT tools, and created new tools specific to each data type. User testing has shown that these additions either drastically reduced the time required to conduct common analysis tasks, or provided the user with analysis options that they did not have before.

I. Introduction

The efficient analysis of crime has become an essential part of current day policing. In a time of rising criminal activity and limited police resources, the ability to conduct practical, computer-based analysis is critical in the prevention of crime. WebCAT is designed to provide straightforward, easy to use analysis tools for users with a wide range of analysis experience. This user-friendly design is critical, considering almost 90% of crime analysts use either no software, or programs such as Microsoft Excel to analyze their data [1].

This project's goal was to improve the accuracy, efficiency, and effectiveness of the WebCAT program. To do this, the project's objectives were to (1) increase the availability and accessibility of crime data and (2) reduce the time required for crime analysis.

In order to increase the availability and accessibility of crime data, the following work was completed:

1. Addition of Querying Capability
2. Adaptation of previous WebCAT tools
3. Creation of new analysis tools

With these additions, users are able to access new data types, and are presented with data type-specific reporting, data viewing, and graphing tools. Users are then able to flow freely between data types, allowing for timely analysis of in-depth data.

The authors conducted usability testing of WebCAT to measure the changes in the availability of data, and the required time for analysis. Testing showed that users were able to complete tasks significantly faster, and in many cases performed analysis they could not have before.

The remaining sections discuss the design and creation of additional data type analysis tools.

Eric J. Ast is with the University of Virginia, Charlottesville, VA 22904 USA (e-mail: eja3d@virginia.edu).

Joshua E. W. Mines is with the University of Virginia, Charlottesville, VA 22904 USA (e-mail: jwm9f@virginia.edu).

Amrita Mukhopadhyay is with the University of Virginia, Charlottesville, VA 22904 USA (e-mail: am2bz@virginia.edu).

Donald E. Brown is with the University of Virginia, Charlottesville, VA 22904 USA (e-mail: deb@virginia.edu).

James H. Conklin is with the University of Virginia, Charlottesville, VA 22904 USA (e-mail: conklin@virginia.edu).

II. Added Querying Capability

The creation of additional querying capability was the first step in allowing WebCAT users to exploit additional data types. For each new data type, we had to decide which parameters to include in the query. An all-inclusive query, featuring each attribute present in the database, a semi-inclusive query which provided only select attributes, and no query were the alternatives evaluated for each new data type.

In selecting alternatives, the speed and ease of use and the presence of desired parameters were the main indices of performance. There exists a tradeoff between these two measures. The inclusion of one additional parameter would increase the presence of a desired field, but how practical that field is to the user, and at what expense its presence is to the query page had to be considered. The aim of the researchers was to include parameters which satisfied the largest proportion of users' needs, while maintaining a usable interface. An all-inclusive query would provide the user all desired parameters, but possibly at the expense of usability. Recognizing the point at which the benefit of one additional parameter began to decrease was essential in the selection process. Following is a breakdown of additional data type querying capability by alternative chosen.

1. No Query

For the offense data type, no additional query was created. This decision was made based on the close relationship between the offense and incident data types. Since WebCAT already possessed comprehensive incident analysis tools, the addition of an offense query would not have increased the program's utility.

2. Semi-Inclusive Query

Semi-inclusive queries were created for the call for service, arrest, person and warrant data types. Taking into

account the large number of attributes present in jurisdictional databases, all-inclusive queries would have been very difficult for the user to navigate without adding significant functionality beyond the utilization of select parameters.

In order to narrow the list of practical parameters, the researchers worked with users and domain experts to solicit the most useful parameters. In addition, the researchers worked through crime scenarios and performed analysis tasks to find which parameters were utilized most often. For each data type, the most useful parameters began to surface with constant use, while others were rarely, if ever utilized. For example as we worked through scenarios with the person data type, parameters pertaining to physical characteristics and past criminal record, such as height, weight, tattoos, and past arrests were used frequently. Other parameters, such as medical information and demeanor were not as practical to query by. This differentiation was made for each data type, and the most commonly used parameters were included in the query page. The final number of parameters present for each data type is shown in table 1.

	All-Inclusive	Semi-Inclusive
Call For Service	15	13
Arrest	28	15
Person	35	20
Warrant	118	11

Table 1: Number of parameters included in each query page alternative [Created by Author].

By creating select, data type-specific query pages, users are able to more efficiently access the data of their choice.

III. ADAPTATION OF VISUALIZATION AND DISPLAYS

Charts, maps, and other visual displays of data provide users the ability to analyze the data more quickly and effectively. Thus, WebCAT must allow analysts to display results specific to each additional data type and in the appropriate form.

Prior to the 2006/07 WebCAT team, a number of graphical displays already existed for crime data. Table 2 lists the charts and mapping options previously provided.

i. 90-Day Chart	ii. Weekly Chart
iii. Weekday Chart	iv. Time Chart
v. Pin Map	vi. Choropleth Map

Table 2: Previous WebCAT visual displays [Created by Author].

With the added functionality of new data types, the existing charts were re-evaluated to most effectively display the new data. The benefit of each individual chart was evaluated based on the nature of data type being analyzed, and the information communicated to the data by the display.

Table 3 shows the final charts and maps that were created for each data type.

	Warrants	Calls for Service	Arrests	People
Charts				
90-Day		X	X	
Weekly		X	X	
Weekday		X	X	
Time		X	X	
Maps				
Choropleth		X		
Pin				

Table 3: Newly developed charts and maps [Created by Author].

For each data type, the chart which most appropriately displayed the data was chosen. Following is a description of each chosen chart.

i. 90- Day Chart

Line Chart: The number of arrests occurring each week would be represented by a data point corresponding to that particular week. A line connecting each successive data point shows the trend in data [2].

ii. Weekly Chart

Line Chart: The number of arrests occurring each week would be represented by a data point corresponding to that particular week. A line connecting each successive data point shows the trend in data.

iii. Weekday Chart

Column Bar Chart: This chart shows the number of arrests occurring per day in the form of a histogram. Each bar would represent one day in the specified time interval.

iv. Time Chart

Circular Line Chart: 4 quadrants denote every 3 hours within a 12 hour period, and a line represents the occurrence of arrests during a particular time frame.

v. Mapping

Choropleth: This map identifies each county within a specified region as a separate entity thus allowing the user to easily interpret the output results.

IV. NEW ANALYSIS CAPABILITIES

With the inclusion of new data types, WebCAT now offers three additional analysis tools. Realizing that manipulating the new data types in the most effective and efficient manner is where the true value in crime analysis lies, WebCAT allows crime analysts to view the data and generate custom reports in ways that convey the most information. The most significant contributions towards making the data "speak" were made in three areas: (1) data linking, (2) monthly reporting, and (3) details views.

A. *Data Linking*

Data linking involves linking unrelated data types by common threads so that a user can move from one data type to another quickly. Data Linking improves crime analysts' efficiency and accuracy. "Using data more effectively allows police to do their jobs better in protecting and defending the citizenry. Reducing the time needed to draw links between crimes allows them to do this" [3].

Because one of WebCAT's top goals is to improve the effectiveness, efficiency, and accuracy of crime analysis, WebCAT offers analysts the ability to maneuver through data very quickly. Consider this process: an analyst queries for arrests made this month. While looking over the data, the analyst notices a Personal Identification Number (PID) that keeps recurring as the arrestee. The analyst decides to take a look at this person's rap sheet. First, the analyst must write down the PID, log out of this query, start another one, and enter the PID in order to get a rap sheet. This process, depending on the system could take one to two minutes. With WebCAT, however, the information is only one click away. WebCAT has links to all information that has hidden details (warrants, people, arrests, calls for service, incidents, offense, etc.). So instead of spending up to two minutes re-querying the system to get a rap sheet of a particular person, the analyst can spend three seconds waiting for a new window to load after he or she clicks on the person's PID from the data they were previously examining. Regardless of the type of data, WebCAT users can navigate quickly across data types.

Figure 1 shows a warrant data view displaying links to other information concerning warrants.

Figure 1: Warrants data view [Created by Author].

B. *Monthly Reports*

During the 2006 DaPro Fall Users' Group Conference, we interviewed crime analysts, and find out what were the key things they wanted from WebCAT. The ability to track their progress in terms of warrants and calls for service were issues that arose multiple times [4]. Therefore, WebCAT offers users the ability to create monthly reports easily.

i. *Alternatives*

WebCAT's reports are designed based off of the COMPSTAT process that the Virginia Beach Police Department uses to show their progress. The statistics that are computed and the charts that are displayed give WebCAT users an easy-to-use, but powerful tool. During the development phase, WebCAT had two alternatives with which it could implement this requirement: (1) manually-designed reports or (2) use Crystal Reports, a business intelligence software package. Table 4 below summarizes these alternatives.

Criteria	Manually-Designed Reports	Crystal Reports
Time	~ 20 Hours	~ 20 Hours
Cost	$0	~ $7,500
Quality	Good	Very Good

Table 4: Monthly Report Alternatives

ii. *Analysis*

Time, cost, and quality were the metrics used to evaluate which of the alternatives was most appropriate. The time required to configure Crystal Reports to work with the WebCAT server was not significantly different from the development time required by the manually-designed reports. While Crystal Reports offered a better quality report due to the flexibility of the parameters, the quality was not worth the $7,500 price tag associated with installing Crystal Reports. To increase flexibility of the manually-designed reports, WebCAT offers a report for each data type in the system and gives users the ability to generate reports for any month of any year and any crime or call type. Giving analysts these options sufficiently meet their needs and surpass their requests. Although Crystal Reports offered a more thorough package, the manually-designed reports more than adequately met the system requirements at a fraction of the cost.

iii. *Implementation*

The following statistics are generated in a monthly report for each data type: (1) the sum for this month, (2) the sum for last month, and (3) the sum for this month last year. Users can choose from one of the following types of reports: (1) incidents, (2) warrants, (3) calls for service, or (4) arrests. For each data type, a basic set of parameters are available which allow users to generate their desired reports.

In addition to the summary statistics, the monthly reports display a 90-Day Chart that shows how many instances of the specified group of parameters occurred on each of the past 90 days. In order to improve efficiency, the chart conveniently highlights the statistically significant days (more than 2 standard deviations away from the mean) for the user. These innovative features combine to offer analysts a high quality report. An example, showing the monthly report for Aggravated Assaults for August 2004, is shown below in Figure 2.

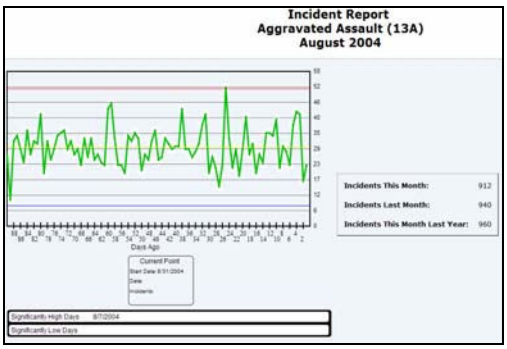

Figure 2: Incident monthly report [Created by Author].

C. Details View

i. Rationale

The details view is an extension of the data linking already described. Its primary purpose is to allow the user to quickly and conveniently get a large picture of the data in a useful manner. Used in conjunction with the data linking, the detail view allows users to get a very detailed view of a particular record in the database. Whether it is a person's rap sheet, which will display all of his/her information from name to social security number to how many and what warrants he/she has out to how many tattoos and where they are located. The details view shows the most detailed data stored in the database.

ii. Implementation

The page design is innovative and flexible allowing users to see the data and developers to modify the data displayed easily. The design is also conducive for reproduction on the analysts' end. Whether the analyst wants to send a link to his/her boss or print it out as a hardcopy, either option is only a few clicks away and take very little time to accomplish. As with the other views in WebCAT, this view offers data linking to other data types. Using the example above, a person's detail view contains links to the details of that person's arrests and warrants. Figure 3 below illustrates a person detail view.

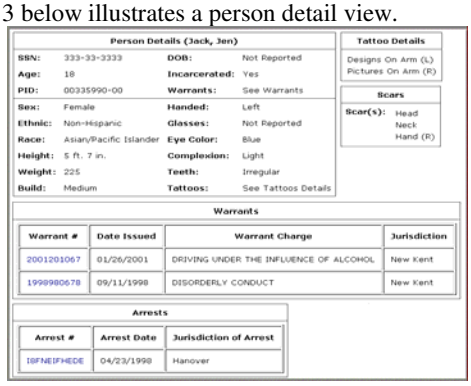

Figure 3: Person detail view [Created by Author].

V. PERFORMANCE EVALUATION

WebCAT was evaluated with a ten subject usability test. Subjects were given a brief introduction to the program, and were then asked to complete seven distinct tasks utilizing WebCAT's new analysis features. The amount of time required for each task, and any additional significant user interaction with the program was recorded and the users were asked to complete the tasks again. The goal of this test was to assess the following five features of the system.

1. Learnability
2. Memorability
3. Efficiency
4. Errors
5. Satisfaction

The manner in which users progressed from the first to the seventh task was a measure of the learnability of the system. While tasks significantly increased in complexity within a data type, users did not require significantly more time to complete them. The first three tasks were all based on the warrant data type. The first task took users an average of 75 seconds to complete and only involved one step. Tasks two and three each involved four tasks, while users only required an additional thirteen seconds on average, completing both the second and third tasks in 88 seconds. After completing only one simple task, users were able to complete more complex tasks in a comparable time period.

The memorability of the system was evaluated by comparing the first time the users completed the seven tasks to the second time they completed the tasks. With ten users completing seven tasks each there were a total of 70 first-run tasks completed. In 69 of these cases, users completed the task more quickly the second time. On average, users required 95 seconds to complete tasks on their first attempt. The second time they completed the tasks, users required only 50 seconds per task. By completing a series of tasks only once, users were able to reduce the time needed to conduct analysis by 47%.

The efficiency, satisfaction, and errors evaluation of WebCAT were all made by observing the user interaction with the system, and by surveying the user after the completion of the test. While time used to complete tasks was an indicator of system efficiency, the manner in which the user progressed through tasks and how much time they wasted with erroneous activity was the observable indication of efficiency we used.

Measurement of system errors was based on user feedback, and examining commonplace mistakes that were made. A common user error was not selecting multiple instances of a parameter, such as outstanding and recalled warrants. The system requires holding of the Ctrl key to multiple select, a feature many users were not aware of. Simple observations like these yielded many minor system errors.

Finally, assessment of user satisfaction was predominantly derived from final user feedback. While observing the users completing tasks yielded some sense of how satisfying the system was, user surveys were the most valuable tool. In general, users indicated that they were satisfied with the system, and felt that the design was intuitive. Sources of dissatisfaction ranged from the confusing listing to certain parameters, to the lack of consistency between certain query pages.

Overall, users were satisfied with WebCAT. User testing yielded several minor errors and showed that the system is user-friendly, learnable, and memorable.

CONCLUSION

This paper describes WebCAT, a system designed to improve criminal incident data sharing and data analysis, within Virginia. This year, the major contributions made towards the project were the incorporation of additional data types and development of tools to specifically analyze the additional data types. Extensive testing and evaluation have determined WebCAT to be a comprehensive analysis system that allows the user to conduct queries with greater efficiency and increase the capabilities of analysis.

ACKNOWLEDGMENTS

The 2006/07 WebCAT team wishes to acknowledge the contributions of the previous five WebCAT teams without whom this project would not be at its current state. We would also like to thank Mr. Mike Lawton of DaPro Systems, Det. Janice Coles of the UVA Police Department and Mr. Butch Johnstone of the VDCJS for providing us with direction towards achieving our goals.

REFERENCES

[1] Almazlinos, C., Bowman, D., Brown, D., Conklin, J. Eagan, R., Kuklinski, T., Nguyen, D., & Hansen, P. (2006). WebCAT: The Development, Performance Analysis, And Deployment of a Web-based Crime Analysis Toolkit. In *Proceedings of the 2006 Systems and Information Engineering Design Symposium*. Charlottesville, VA: University of Virginia Department of System and Information Engineering.

[2] Wikipedia (2006). *Line Chart. Retrieved April 3. 2007 from:* http://en.wikipedia.org/wiki/Line_chart.

[3] Gilad, Benjamin & Tamar Gilad. (1988). *The Business Intelligence System: A New Tool for Competitive Advantage*. American Management Association: New York, NY.

[4] Crime Analysts from DaPro Fall Users' Conference. Personal Contact. October 12-13, 2006. Virginia Beach, Virginia.