The author(s) shown below used Federal funds provided by the U.S. Department of Justice and prepared the following final report:

Document Title: **Artificial Neural Network System for Classification of Offenders in Murder and Rape Cases, Executive Summary**

Author(s): **Lars Kangas**

Document No.: **190983**

Date Received: **11/29/2001**

Award Number: **97-IJ-CX-K007**

This report has not been published by the U.S. Department of Justice. To provide better customer service, NCJRS has made this Federally-funded grant final report available electronically in addition to traditional paper copies.

Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

190983

Artificial Neural Network System for Classification of Offenders in Murder and Rape Cases

Summary

Prepared by Lars Kangas for the
National Institute of Justice,
Office of Science and Technology
Under Award Number 97-IJ-CX-K007

When a serial offender strikes, it usually means that the investigation is unprecedented for that police agency. The volume of incoming leads and pieces of information in the case(s) can be overwhelming as evidenced by the thousands of leads gathered in the Ted Bundy Murders, Atlanta Child Murders, and the Green River Murders. Serial cases can be long-term investigations in which the suspect remains unknown and continues to perpetrate crimes. With state and local murder investigative systems beginning to crop up, it will become important to manage that information in a timely and efficient way by developing computer programs to assist in that task. One vital function will be to compare violent crime cases from different jurisdictions so investigators can approach the investigation knowing that similar cases exist.

The "Artificial Neural Network System for Classification of Offenders in Murder and Rape Cases" project developed two software prototypes that demonstrate developed algorithms for analyzing and comparing large databases of crime data. The CATCH (Computer Aided Tracking and Characterization of Homicides) and CATCHRAPE software applications analyses homicide data and sexual assault data, respectively. Both applications are similar, although they analyze different databases. "CATCH" will from hereon refer to both applications.

CATCH was developed to provide crime analysts enhanced means for interpreting large databases of crime data. These databases store a large number of crimes with each case described in a large number of details. The huge volumes of information collected for crimes during investigations make it difficult to compare a crime against other crimes by the investigators alone. CATCH using advanced algorithms running on computers facilitates the investigators need to compare crime data and to find crimes that have similar patterns.

The development of CATCH was made possible with the HITS (Homicide Investigation Tracking System) database system. Police involved in the infamous Green River and Ted Bundy murder investigations in the State of Washington developed HITS more than ten years ago to enable computer-based analysis of murders. The database now contains several thousand violent crimes primarily from the Pacific Northwest. The HITS database system is similar in the crime details it stores to the VICAP system at FBI and the violent crime database system operated by the RCMP in Canada.

The algorithms in CATCH learn from crime data. HITS provided the required historical crime data for CATCH to learn offender behavior so that new unsolved crimes can be characterized for analysis and comparison.

PROPERTY OF
National Criminal Justice Reference Service (NCJRS)
Box 6000
Rockville, MD 20849-6000

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

The CATCH software assists with crime investigations by assessing likely characteristics of unknown offenders, by relating a specific crime case to other cases, and by providing a tool for clustering similar cases that may be attributed to the same offenders. An investigator working with CATCH is working against the collection of crimes in the HITS database. The software speeds up the database access to these crimes, improves the views of the crimes, and organizes the crimes according similarity.

CATCH is a collection of tools that assist the crime analyst in the investigation process by providing advanced data mining and visualization capabilities. These tools include clustering maps, query tools, geographic maps, timelines, etc. Each tool is designed to give the crime analyst a different view of the case data.

The clustering tools are based on artificial neural networks (ANNs). The ANNs in CATCH learns from a subset of existing crime cases (approx. 5000 homicides, 3000 sexual assaults) in the HITS database using parameters describing modus operandi (MO) and signature characteristics of the offenders. The ANN clusters cases onto two-dimensional visual representations, according to the similarity of offender patterns. The proximity of cases within the feature map allows the analyst to identify serial homicides/sexual assaults.

A case being investigated is compared to the solved adjacent cases in the visual representation of the database. A set of one or more cases is analyzed in detail to determine whether a linkage exists between the cases. For example, is it one offender that is responsible for a series of crimes? The investigator goes between the different tools in CATCH and analyses when and where the crimes were committed. The analyst can focus on specific details that he or she believes are important for determining whether there is a connection between the crimes. The analysis work continues using a set of cases where individual cases can be dropped or new cases can be added from the database. All accesses to the database are fast and simple using point and click mouse methodology.

Access to a database system, like the HITS database, requires strict adherence to a query language to retrieve information. This language is slow and difficult to use even for an experienced database professional. SQL is the most common language for accessing a database. CATCH removes the need for an investigator to know this language by dynamically creating the queries in the SQL language on demand, as the investigator accesses the database. The investigator can concentrate on the information in the databases, rather than the access method. The investigator only sees the information in CATCH, not the query language, as in the previous version of software accessing the HITS database.

CATCH characterizes an unknown offender by comparing his or her crime to known offenders of other crimes where the behaviors of the offenders have been similar. The characterization can give some clues to personal information about the unknown offender to expedite the investigation process. This information include, level of education, type of work, physical appearances such as neatness, misuse of alcohol and drugs, etc. The underlying assumption is that two individuals that otherwise behave similarly, may also commit crimes with similar

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

behavior. The validity of this assumption has been proven to some degree by FBI's behavioral science unit and others.

The characteristics of offenders developed in CATCH are hints at best and differs from most of the other information, which are factual and retrieved from the database. With the exception for the characterization tool, CATCH was not designed to modify or generate data describing the crimes in the database. It only organizes the existing facts to provide more meaning to the investigators.

CATCH was developed using visionary information technologies to advance crime analysis. These technologies include methods for analyzing data, visualizing information, and accessing data. Research staff drew from past experience in these technologies and pushed the envelope just a little further by combining these technologies in one application. The application has received considerable attention in the media as being new and innovative.

The investigations of crimes involve analysis of offender behavior. A set of automated technologies can make the work of investigators more efficient, they cannot replace the investigators. Although technologies are becoming more advanced, a software tool should still only be in support of the human in-depth understanding of behavior. The quintessential objective in the research was to facilitate the work of investigators.

The development of CATCH would not have been possible without the collaboration Battelle enjoyed with Washington State Attorney General's Office, Criminal Division. The experience of the crime analysts and the extensive experience they have in operating a database of violent crimes were invaluable to the project. These crime analysts were the drivers that recognized the need to research the possibility of using advanced technologies to make their work more efficient.

CATCH demonstrates the value of incorporating new technologies and methods that can aid the work in crime investigations. The scope of the project was limited to specifically show how the artificial neural network technology can support a multivariate or information fusion approach to finding significant patterns in crime data. Other research and development efforts concentrate on equally important aspects of crime analysis, such as the use of graphical information systems. Although CATCH was developed as a self-sufficient application and includes some crime mapping and other tools of a complete application for crime analysts, it did not explore or exhaust the possibilities of those tools.

Two versions of CATCH were developed, CATCH and CATCHRAPE for homicide data and sexual assault data, respectively. These applications were deployed at Washington State Attorney General's Office, Criminal Division for evaluation. Although the first versions of CATCH were configured specifically for the HITS database of violent crimes, they can be applied against other crime databases through relatively minor changes in the software.

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

Status

PNNL finished Version 1.0 of CATCH (Computer Aided Tracking and Characterization of Homicides). A second application CATCHRAPE to analyze sexual assaults was also completed with a subset of the functionality of CATCH for murders.

A parallel project at PNNL, sponsored by Washington State, redesigned the HITS database after the completion of CATCH, Version 1.0. A development of CATCH, Version 1.1 was started to accommodate the new HITS database. This version is operational at PNNL using an estimated 95% of the crime details from the new version of the HITS database. (The decision to redesign the HITS database came very late in the CATCH development, when most of the project time and funds had been spent. Version 1.1 has not been completed for installation at Washington State AGO to run against the current HITS database).

Washington State AGO started an evaluation of CATCH. The evaluation was not extensive due to loss or availability of key staff at AGO, which included Dr. Robert Keppel's retirement and Ken Hanflen's departure.

Preliminary evaluations by Dr. Keppel, et. al., suggest that the clustering algorithms and visualization tools in CATCH have the potential to add considerable value to crime analysts.

Presentation of CATCH to NIJ

PNNL is scheduled to present CATCH at an "In Process Review" arranged by Mr. William Deck, NLECTC Southeast to Mr. Trent DePersia, NIJ, et.al., May 22, 2000 in Denver, Colorado.

Description of CATCH

CATCH was developed to assist crime investigations by assessing likely characteristics of unknown offenders, by relating a specific crime case to other cases, and by providing a tool for clustering similar cases that may be attributed to the same offenders.

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

CATCH is a collection of tools that assist the crime analyst in the investigation process by providing advanced data mining and visualization capabilities. These tools include clustering maps, query tools, geographic maps, timelines, etc. Each tool is designed to give the crime analyst a different view of the case data.

The clustering tools in CATCH are based on artificial neural networks (ANNs). The ANNs learn to cluster similar murder and sexual assaults cases residing in a database. The clustering algorithm is applied to parameters describing modus operandi (MO), signature characteristics of the offenders, and other parameters describing the victim and offender. The proximity of cases within a two-dimensional representation of the clusters allows the analyst to identify similar or serial murders and sexual assaults.

Future Work

We finished a first draft proposing a new version of CATCH, which incorporates additional tools that have been identified from the current research and development. The first set of tools in Version 1 of CATCH was concentrated on researching the value of using artificial neural networks to cluster similar cases. The new tools will provide the crime analysts a more complete suite of tools.

Presentations of CATCH

- L.J. Kangas, K.M. Terrones, R.D. Keppel, R.D. La Moria. 1999. "Computer Aided Tracking and Characterization of Homicides & Sexual Assaults (CATCH)." SPIE AeroSense '99, Orlando, FL, April 5-9.
- L.J. Kangas. 1999. "CATCH." Justice Research and Statistics Association, Washngton, DC, September 30.
- L.J. Kangas, K.M. Terrones. 1999. "Data Mining and Visualization for Crime Analysis: CATCH." Presented at 8th Advanced Information Processing & Analysis Steering Group, Washington DC, March 23-24.
- S.V. Shoemaker, L.J. Kangas, 1998. "Washington Courts Database Conversions." Information Sharing Technology Workshop, Georgia Institute of Technology, May 6.

Media Exposure of CATCH

- "Computer Takes a Byte out of Crime." Popular Mechanics, March 2000.
- "CATCH Program." TECH beat, Fall 1999.
- "New Way to CATCH Bad Guys." http://www.govtech.net/onlineservices/news/newspage.shtm?story=June181999NewWaytoCATCHBadGuys, June 18, 1999.
- "Catch as CATCH Can." MIT's Magazine of Innovation Technology Review, March/April 1999
- "Buying Computers and Software" (CATCH software described) Law Enforcement Technology, April 1999
- "CATCHing Criminals." http://newsweek.com/nw-srv/issue/15_99a/tnw/today/cs/cs01th_1.htm, April 9, 1999
- "Thinking Like a Serial Killer." http://www.wired.com/news/print_version/technology/story/19940.html?wnpg=all, June 1, 1999
- "David Lawrence talking with Lars Kangas, Sr. Research Engineer at Pacific Northwest National Labs" http://www.audiocast.net/geekradio, June 7, 1999.

PROPERTY OF
National Criminal Justice Reference Service (NCJRS)
Box 6000
Rockville, MD 20849-6000

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.