

Combating crime in Gauteng, South Africa: a paradigm shift

C. E. Cloete & J. S. Spies

*Department of Construction Economics, University of Pretoria,
South Africa*

Abstract

South Africa offers a valuable laboratory to study criminality because of its high levels of crime due to a heterogeneous environment of people, cultures, and economic development.

This paper assesses the limitations of conventional practices in combating crime and illustrates the application of innovative artificial intelligence software to assumed unrelated databases – demographic, geographic, socio-economic and others – to find more effective ways to prevent criminal incidents. True unsupervised machine learning artificial intelligence software developed in South Africa provides neural network prediction of criminal incidents by identification of supposedly unrelated variables.

Application of the software to find the hidden critical factors leading to bank robberies is illustrated in a specific case, where a robbery at a specific bank branch was predicted with an exceptionally high level of accuracy.

Keywords: combating crime, artificial intelligence, causal layering analysis, neural networks, bank robberies, crime prediction.

1 Introduction

South Africa offers a valuable laboratory in which to study crime-combating measures due to its high levels of crime in a society of different cultures, widely disparate levels of economic development and education, and uncontrolled migration of foreign migrants across its borders and to urban areas.

This paper illustrates the use of an Artificial Intelligence programme (RapAnalyst) as a significant enabler to shift from reactive crime prevention to proactive strategies and initiatives. It also argues for the establishment of so-



called *Fusion Platforms* for breaking down data silos and enabling optimal synergy of cutting-edge technology and innovation to counter crime.

A specific case study will be presented and discussed to illustrate the essence of these paradigm shifts.

2 The crime levels in Gauteng

Of the nine provinces in South Africa, Gauteng is by far the smallest (figure 1).

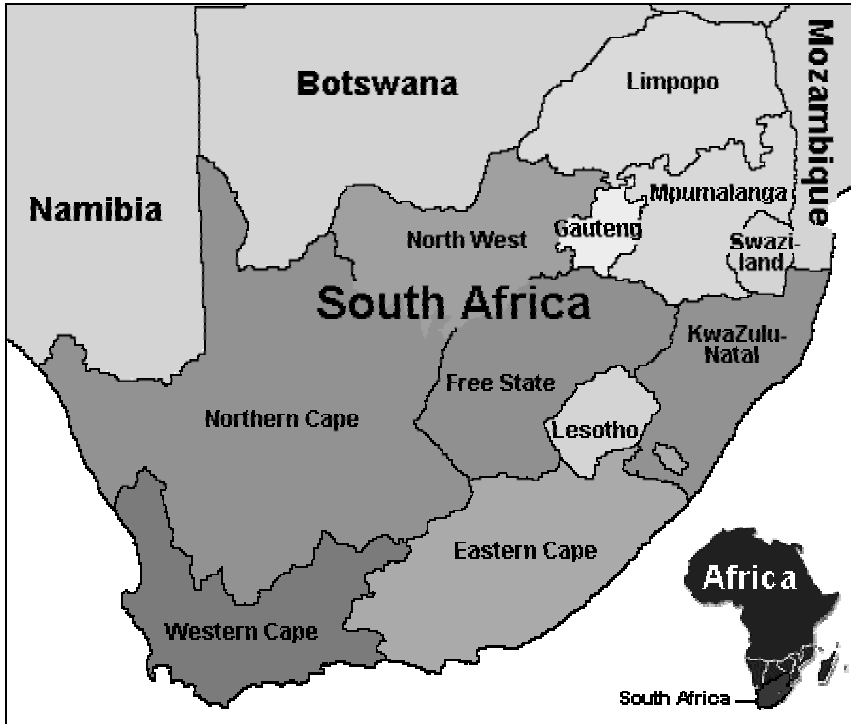


Figure 1: Gauteng province in relation to South Africa.

Nevertheless, it is home to three of the country's important metropolitan areas, namely Johannesburg, Tshwane (Pretoria) and Ekurhuleni (comprising various cities east of Johannesburg). While Gauteng comprises only 1.4 percent of South Africa's surface area, it is home to some 21.5 percent of the population, supplies work to almost 30 percent of the labour force and accounts for virtually half of the private sector building activity. The three metropolitan areas in Gauteng are highly industrialised, with a well-developed transport system for road and rail traffic, including extensive urban freeway systems, metropolitan bus networks and a rail network for freight and passenger transport.

Given the above, it is not surprising that Gauteng is the province with the highest crime statistics.

The 15-year period after the 1994 peaceful transfer of power to a democratic government also led to an exponential rise in all crime categories of:

- *White Collar Crimes* that are mostly boundary-less, faceless and global, of which some types require zero cycle-time to counter, e.g. fraud, embezzlement, money laundering, computer hacking.
- *Crimes against Property* targeting physical assets at a geographic place or in transit, e.g. theft, vandalism.
- *Crimes against Persons*, entailing physical attacks, attacks on personal rights, entitlements, or human values.

Many white-collar crimes are especially difficult to prosecute because the perpetrators are sophisticated criminals who conceal their activities through a series of complex transactions and operating in the e-world. Yet government and private sector continue to ring-fence their own operating space, and persist in combating crime in a 'reactive' mode within these demarcated boundaries; while criminals and criminal syndicates, on the other hand, are not limited by these boundaries and are in fact capitalising on this fragmentations and attitudes.

3 The shift from reactive to proactive sharing of data

A criminal event is only a manifestation of a single incident in a larger criminal process of activities that need to be understood. Dealing reactively with crime as events can at best only enable incremental and limited countering improvements against crime. Understanding the crime processes and finding the hidden factors that trigger criminal activities will be the only enabler to find more innovative solutions to proactively counter criminal intent.

The authorities, professional disciplines, communities, and individuals must unite and act in the present tense, strip down information and data boundaries and start applying leading technologies that can make a real difference to protect the tomorrows, instead of discussing tomorrow how to deal with yesterday's criminal incident.

Today's world is inundated with data and huge databases that are normally only applied for one purpose before being archived. Immense breakthrough possibilities exist by linking the data that can be applied in combating crime, which is archived by the Police Services, the Metropolitan Councils, security companies, NGOs, universities, Internet data resources, research institutions, etc. These sources that can be applied synergistically in combating crime in a proactive manner. By applying intelligent machine learning software that is self-learning this available data, which is quite often complex and multi-dimensional, can be analysed to uncover hidden relationships.

This has become reality with intelligent software developed in South Africa that now can be applied on data from different databases to find hidden relationships between different seemingly unrelated data attributes, distinguish between strong relationships and weak relationships, and enabling predictions of extraordinary accuracy. [Note that 'intelligent' used in this context is a self-learning neural network ability to draw conclusions based on certain findings, and does not imply 'creative' thinking that is unique to man that can 'dream' and 'imagine'.]



4 Artificial Intelligence as a significant enabler to shift from reactive to proactive strategies and initiatives

4.1 Artificial Intelligence

Artificial Intelligence (AI) software is defined for the purpose of this paper as “software that transforms complex data into actionable information without any prior assumptions and rules.” AI that can interactively work with the analyst, which also enables understanding of complex multi-dimensional data sets through clear visualization (Bishop [1]; Mitchell [5]).

Unlike traditional statistical techniques, AI must deal with data holistically and shows the relationships between all attributes in data sets, as opposed to targeting specific attributes that are assumed to be relevant. This allows AI to deliver high levels of predictive relevance, even when dealing with missing data or data that has been poorly measured (Duda et al. [2]; Engelbrecht [3]; MacKay [4]; Negnevitsky [7]) and has led to application in a number of areas (e.g. McDuling [6]).

Self-learning AI computer software leverages the biological brain functions to comprehend complex multi-dimensional data that is not based on predefined rules, prejudice and personal assumptions (Weiss and Kulikowsky [9]).

In the following case study the AI software employed was RapAnalyst. (RapAnalyst was developed by C. Wocke and R. Britz in South Africa during the 1990s, and all IP rights were bought by BMGI in the USA).

4.2 A case study

The following case study is a pilot project that was done during the first years of the new millennium on bank hall robberies and other bank crimes in Gauteng.

4.2.1 The scenario

After 1994 the bank halls in South Africa was targeted by organized bank robbery and other bank crime syndicates. These bank robberies were true Hollywood style with guns and fast getaway cars, leaving traumatized bank employees and citizens in its wake.

The number of incidents grew dramatically until 1999, when the Banking Council of South Africa convinced all its member banks to unite against this scourge and a private local consulting house was contracted to research and propose more effective crime combating strategies.

This led to several strategies to combat violent bank crimes, one strategy of which was to establish a central database on which all violent criminal incidents were to be recorded within 48 hours of an incident. This database recorded significant more details about each criminal incident, in contrast to the South African Police Service that only recorded higher level definitions of crime e.g. ‘housebreaking’ that covered everything from domestic burglaries, bank robberies, etc.

This more detailed data enabled the identification of preferred bank robbery areas and bank groups to be targeted. Using traditional statistical methods it



become possible to identify some robbery patterns and predict so called ‘hot spots’ for visibly policing. This, with support of other combating strategies that were implemented in collaboration with the South African Police Service broke the exponential growth in bank robberies. Sharing banking violent crime database information with the Police, special police operations and projects were put in place that turned the tide significantly. Bank robberies then started to decline as more perpetrators were arrested and incarcerated.

During 2004 the existence of RapAnalyst came to the attention of a database analyst at one of the banks and he requested the supplier of the software to engage in a pilot project. The outcome of the project was a huge success in building ‘knowledge filters’ that identified the specific risk factors to be attended to for each bank branch to increase its resilience against bank robberies. RapAnalyst also demonstrated amazing prediction ability on every intervention taken. Unfortunately, the security management on violent crime of that time rejected the project, describing it as ‘impossible’. However, the success of the programme led to the creation of the ‘Griffin’ suite of services with RapAnalyst in its core to combat bank robberies in the USA.

4.2.2 How does RapAnalyst function?

RapAnalyst software transforms complex data into actionable information using SOM (Self Organising Maps) AI (Artificial Intelligence) methodology delivering an understanding of complex data sets through visualization maps.

Unlike traditional statistical techniques, RapAnalyst approaches data analysis holistically and shows the relationships between ALL attributes in a data set, as opposed to targeting specific attributes that are assumed to be relevant. This allows RapAnalyst to deliver high levels of predictive relevance, even when dealing with missing data or data that has been poorly measured.

“RapAnalyst identifies:

- Relevant attributes within data related to a current problem
- Significant relationships between relevant attributes
- Outcomes of predictive “what-if” scenario modelling

By displaying the interrelationships between data points in an easily interpreted visual format, RapAnalyst allows the user to identify opportunities and make decisions with confidence. The application delivers thorough and accurate data understanding that is not restricted by inherent assumptions.

RapAnalyst allows one to work with high-dimensional data within a two-dimensional visualization called “Knowledge Filters.” A Knowledge Filter is an optimized representation of the input data set that is constructed during the RapAnalyst training process. This training process, which uses unsupervised neural networks, arranges a representation of the data in the Knowledge Filter based on similarity.

The Knowledge Filter can then be applied to investigate every attribute in the data set using the resulting attribute windows. It is when several different attribute windows are compared that the investigative power of RapAnalyst emerges.

The attribute window views are a colour depiction of complex multi-dimensional data in two dimensions – a bird’s eye view of the data. As each



attribute is displayed in its own window, one can easily identify the dynamics and interrelationships within the data. This depiction can also quickly provide insight into how and why certain events occur.

4.2.3 Application of RapAnalyst to bank robbery data

The following data on bank robberies and other bank crimes was extracted over a period of a few years to enable pattern recognitions that could be of seasonal nature:

- Date and time of incident
- Type of crime incident
- Bank branch details
- Location details
- Close proximity factors
- General location factors

These data elements can be seen in more detail in Figure 2 as headers of ‘dropdown boxes’.

The question may be asked why other types of crime against banks like ATM mugging, burglary, card jamming, etc are also included if the analyst is only focussing on bank hall robberies? The answer is that although these crimes may seem unrelated, RapAnalyst may find hidden relationships that could be significant and add value to the analyses. This, specifically, is one of the many benefits in using unprejudiced machine learning software.

In the specific case, took RapAnalyst a few minutes to analyse the given data into knowledge filters that identified and made visible 2-dimensional maps of all significant data clusters, indicating strong clustering relationships e.g. the different crime types, bank branch data that are very prone to bank robberies and also that repel bank robberies, and many other relationships.

Visual 2-dimensional pictures in colour (adding a 3rd visual dimension) with an interactive vertical scale at the bottom of the specific data attribute map with values between its minimum and maximum boundaries, are given for each data variable or non-numerical descriptors. The non-numerical descriptors are binarised to a 0 or 1 to depict a “no” or “yes” for the specific descriptor.

The analyst now interprets these maps but can also interact with them by clicking on specific icons on a toolbar and then any place in a map to get clarity on different relationships and interactivity between the different data elements and its attributes, He can actually have an interactive ‘discussion’ using the ‘what if?’ functionality with the ‘Knowledge Filter’ e.g. ‘what if I replace the entrance door from the current type to the other type where branch robberies is low, and shift my parking area in accordance?’ RapAnalyst will respond immediately with the calculated change in the risk level of a future bank robbery.

These interactive analyses are obvious of immense value in outcome for improved decision taking on how to combat crime; but also to optimise the placing and design of future bank branches to reduce the risk of bank hall robberies. The template in seen in Figure 2 was designed to do exactly this.

However, when the branch is already in existence the same template can be used to assess the specific bank branch’s risk level by selecting specific entries within each of the dropdown boxes that is applicable to the specific branch; and

SABRIC : Violent Crime

Financial Institution Location: <input type="text"/> Distance to freeway ramp: <input type="text" value="4 - 6 km"/> On the way to freeway ramp there are: <input type="text" value="Yield Signs"/> Taxi rank near bank: <input type="text"/> Townships: <input type="text"/> Security complex/secluded village: <input type="text"/>		General Location Area: <input type="text"/> Sabric Area: <input type="text"/> Suburb: <input type="text"/> Social Status: <input type="text"/> Parking area near bank: <input type="text"/> Parking facilities: <input type="text"/> Pedestrian movement: <input type="text"/>																															
Business Area near Bank Are there restaurants/coffee shops/taverns near (radius of 10 meters) the the bank, with a view of the bank/ATM: <input type="text"/> Department stores/Supermarkets: <input type="text"/> General Small Businesses: <input type="text"/> Cash Loans: <input type="text"/> Café: <input type="text"/> Guards at or in vicinity of bank: <input type="text"/> Metro police/car watchers: <input type="text"/> Police Officers: <input type="text"/>		<input type="checkbox"/> Bank is in Shopping Centre Distance to entrance/exit: <input type="text"/> Distance to escalator/lift: <input type="text"/>																															
Crime Incident Types <table border="1"> <tr><td>No crime:</td><td>0%</td><td><input type="text"/></td></tr> <tr><td>ATM Mugging:</td><td>14%</td><td><input type="text"/></td></tr> <tr><td>Attempted Bank Robbery:</td><td>25%</td><td><input type="text"/></td></tr> <tr><td>Attempted Burglary:</td><td>11%</td><td><input type="text"/></td></tr> <tr><td>Bank Robbery:</td><td>75%</td><td><input type="text"/></td></tr> <tr><td>Burglary:</td><td>0%</td><td><input type="text"/></td></tr> <tr><td>Card Jamming:</td><td>0%</td><td><input type="text"/></td></tr> <tr><td>Card Swopping:</td><td>0%</td><td><input type="text"/></td></tr> <tr><td>Cheque Fraud:</td><td>0%</td><td><input type="text"/></td></tr> <tr><td>Cash in Transit:</td><td>0%</td><td><input type="text"/></td></tr> </table>		No crime:	0%	<input type="text"/>	ATM Mugging:	14%	<input type="text"/>	Attempted Bank Robbery:	25%	<input type="text"/>	Attempted Burglary:	11%	<input type="text"/>	Bank Robbery:	75%	<input type="text"/>	Burglary:	0%	<input type="text"/>	Card Jamming:	0%	<input type="text"/>	Card Swopping:	0%	<input type="text"/>	Cheque Fraud:	0%	<input type="text"/>	Cash in Transit:	0%	<input type="text"/>	Bank Entrance System: <input type="text"/> Inside of bank visible from outside: <input type="text"/> ATMs part of bank structure: <input type="text"/> # of ATM's at bank?: <input type="text"/> Guards on duty at ATM?: <input type="text"/> Guards at ATM: <input type="text"/>	
No crime:	0%	<input type="text"/>																															
ATM Mugging:	14%	<input type="text"/>																															
Attempted Bank Robbery:	25%	<input type="text"/>																															
Attempted Burglary:	11%	<input type="text"/>																															
Bank Robbery:	75%	<input type="text"/>																															
Burglary:	0%	<input type="text"/>																															
Card Jamming:	0%	<input type="text"/>																															
Card Swopping:	0%	<input type="text"/>																															
Cheque Fraud:	0%	<input type="text"/>																															
Cash in Transit:	0%	<input type="text"/>																															
		Day of Week: <input type="text" value="Thursday"/> Day of Month: <input type="text" value="5"/> Month of Year: <input type="text" value="March"/> Time: <input type="text" value="07h00-09h"/>																															
		<input type="button" value="Clear"/> <input type="button" value="Predict"/>																															

Figure 2: Template illustrating the attributes used in predicting a bank robbery at Woodlands Boulevard, Pretoria.

can also be used to identify that changes need to be made to reduce the risk to an acceptable level.

Within the dropdown boxes all numerical values or descriptors are recorded that falls within the domains of the different data elements and their attributes. Predictions are then made by selecting the dropdown values and descriptors of what is known of the branch and entering a “?” in the value or descriptor space



you need a predicted value or choice for, before clicking on the ‘Predict’ icon on the bottom right corner of the template. The ‘?’ will then be replaced with the predicted outcomes.

In applying this template to different branches in existence it became evident what aspects in the design features attracted or repelled bank robbers. This is done by entering the true data of the branch and then predict on ‘?’ in the time, month and day boxes. The outcome is interpreted as:

- The ‘Day of Week’ is a preferred day by robbers for that bank branch profile e.g. Thursdays with low bank hall visits;
- The ‘Day of Month’ is indicative of the referred period e.g. just after the end of the month when all business cash has been received
- ‘Month of Year’ may show up as festive periods e.g. December for beach front branches
- A ‘Time’ preference was found to correlate strongly to non-traffic peaks.

Using the ‘what if’ functionality icons of RapAnalyst in the design and the location of new bank branches can easily assist in selecting the better site and adapting the design of the branch to that environment.

This template was applied in 2004 to predict how long it would be after the opening of a new bank branch with a particular design in the newly-developed Woodlands Boulevard Shopping Centre, Pretoria before it became a target of a criminal attack. The unexpected prediction was ‘Within two weeks’ – and the branch was hit before the end of the 2nd week.

These predictions can be further improved by adding data from other sources that could have some possible relationship to these crimes, and letting RapAnalyst determine possible relations. This raises the question on how to break down the ‘silos’ of data held by different organisations. The next section deals with a solution that is currently making a breakthrough in South Africa.

5 Fusion platforms

To enhance the value of AI analyses it is of great importance to break down barriers that block the sharing of data. These barriers can however be overcome by the establishment of “Fusion Platforms”. The Innovation Hub in Pretoria, in which the University of Pretoria also has a vested interest, established a Safety and Security Fusion Platform (S&S FP) to address this particular need.

The S&S FP is a neutral platform for think tank workshops on safety and security issues is facilitated by a neutral professional facilitator. Participants are drawn from:

- Government
- Academia
- Private Sector

The role of the S&S FP is to enable unfettered discussions to find ‘fusion points’ between the associated members, data and technology.

The key is to move the participants outside their traditional comfort zones to focus on breakthrough solutions on how to break the legacy moulds that do not have an impact any longer, and to find new and innovative ways in creating a safe and secure future for Southern Africa.



A example of the creation of a fusion platform is the initiative of the South African banking industry in the late 1990s when bank robberies grew exponentially to just less than one a day. The individual banks could not fight back. For the first time in the South African history all the different banks took hands and decided to share data on violent crimes to combat the bank robberies as a united force. A co-operation agreement was also signed with the South African Police Service and the Banking Council of South Africa to share this data because the police data was not broken down to the level of detail as that which was recorded by the banks. The data analysis, although very basic in the beginning, was of such value that joint police and bank interventions reduced the bank hall robberies by two thirds within four years. This illustrates the value of sharing a platform between the banks and the South African Police Service.

The next phase is developing a clear strategy to enhance the value of the 'Fusion Point' between data and technology to reduce crime levels proactively in Southern Africa – see figure 3.

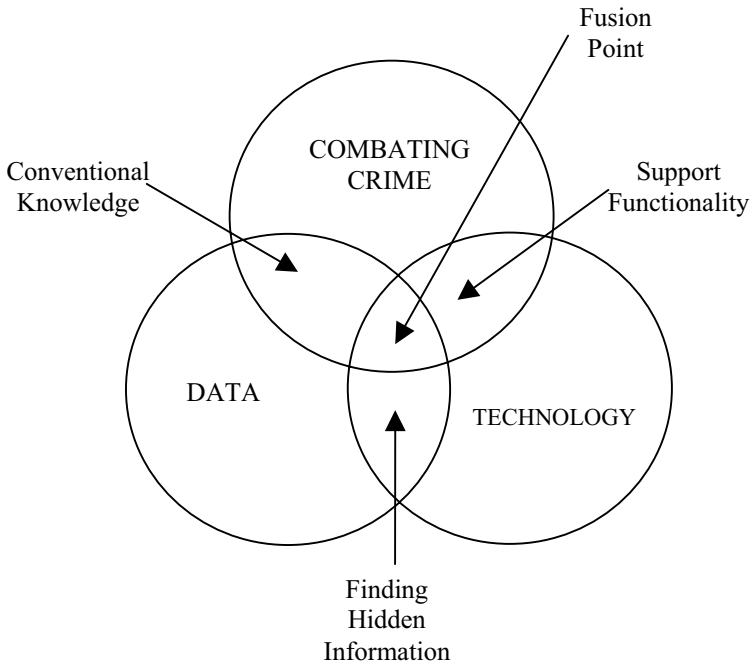


Figure 3: The fusion point between data, technology and combating of crime.

6 Conclusion

It became very evident that for the Safety and Security Fusion Platform to become truly proactive it will be necessary to do foresight studies, and in

particular give attention to Africa's high levels of crime due to a heterogeneous environment of people, cultures, and economic development.

This implies analyses methodologies like CLA (Causal layered analysis). CLA ... "is one of several futures techniques used as a means to inquire into the causes of social phenomena and to generate a set of forecasts as to the future course of the phenomena.

As a theory, CLA seeks to integrate empiricist, interpretive, critical and action learning modes of knowing (loosely, science, social science, philosophy and mythology). As a method, its utility is not in predicting the future but in creating transformative spaces for the creation of alternative futures. It is also likely to be of use in developing more effective — deeper, inclusive, longer term — policy." [http://en.wikipedia.org/wiki/Causal_layered_analysis, accessed 26 March 2009]

According to this source, the advantages of CLA for the Fusion Platform will be that it:

"(1) Expands the range and richness of scenarios; (2) When used in a workshop setting, it leads to the inclusion of different ways of knowing among participants; (3) It appeals to and can be used by a wider range of individuals as it incorporates non-textual and poetic/artistic expression in the futures process; (4) CLA layers participant's positions (conflicting and harmonious ones); (5) It moves the debate/discussion beyond the superficial and obvious to the deeper and marginal; (6) It allows for a range of transformative actions; (7) CLA leads to policy actions that can be informed by alternative layers of analysis; and (8) CLA reinstates the vertical in social analysis, that is, from postmodern relativism to global ethics."

Applying AI to find the hidden relationships between many different data sets in an extremely complex space of multi-dimensions will enable a true understanding of the significant factors that trigger crime and how to counter or mitigate these events in a proactive manner.

References

- [1] Bishop, C. M., *Neural Networks for Pattern Recognition*, Oxford University Press, 1995.
- [2] Duda, Richard O; Hart, Peter E. and Stork, David G., *Pattern classification* (2nd edition), New York: Wiley, 2001.
- [3] Engelbrecht, Andries P., *Computational Intelligence. An Introduction*. Chichester, New Essex: Wiley, 2002.
- [4] MacKay, D. J. C., *Information Theory, Inference, and Learning Algorithms*, Cambridge University Press, 2003.
- [5] Mitchell, T., *Machine Learning*, McGraw Hill, 1997.
- [6] McDuling, J.J., *Towards the Development of Transition Probability Matrices in the Markovian Model for the Predicted Service Life of Buildings*, PhD Thesis, Department of Civil and Biosystems Engineering, Faculty of Engineering, Built Environment and Information Technology, University of Pretoria, South Africa, 2005.



- [7] Negnevitsky, M., *Artificial Intelligence, A Guide to Intelligent Systems*. Essex, England, Addison-Wesley, 2002.
- [8] Russell, Stuart and Norvig, Peter. *Artificial Intelligence. A Modern Approach*. Upper Saddle River, New Jersey: Prentice Hall, Pearson Education International, 2003.
- [9] Weiss, Sholom and Kulikowski, Casimir., *Computer Systems That Learn*, Morgan Kaufmann, 1991.

Additional information on the software programme can be obtained from raptor@raptorinternational.com or www.raptorinternational.com.

