

# Agent-Based Behavioural Modelling for Anomaly Detection in Call Data from Telecommunication Networks

Isaac O. Osunmakinde and Anet Potgieter  
Agents Research Group (ARG)  
Department of Computer Science  
Faculty of Science  
University of Cape Town  
Cape Town, Republic of South Africa  
E-mail: ([segun@cs.uct.ac.za](mailto:segun@cs.uct.ac.za) and [anet@cs.uct.ac.za](mailto:anet@cs.uct.ac.za))

## Abstract

**Anomaly detection in telecommunications data tries to discover deviant behaviour of individual subscribers, including for example: detection of inconsistencies in call data, such as customer churn or attrition, potential fraud, deliberate or unintended expensive mistakes in call data, and so on. These have consequently led to unquantifiable loss of revenue to many telecommunication networks world-wide. Although the intentions of most subscribers to these networks are unknown when making phone calls, their service consumption pattern is reflected in their call data.**

**Recent studies have investigated the challenges of anomaly detection but have not given conclusive solutions to address this problem holistically. In this work, we infer that if appropriate anomaly indicators for individual subscribers are used for detection, the true positive rates of the approaches can be maximized, while the false alarms can be minimized. The challenge addressed in this research is to find a technique of efficiently identifying the indicators that facilitate anomaly detection methods, providing qualitative knowledge. This knowledge will assist analysts and managers in explanatory, exploratory and decision making activities. We address this challenge by using an Agent-Based Behavioural Bayesian Network (BBN). This BBN models an individual subscriber by using an evolutionary algorithm, and anomaly indicators are identified through emergent behaviour. Hence, our implementation results for land-line subscribers provide a general approach of improving the detection qualities of existing anomaly detection methods in telecommunication networks.**

**Keywords: Subscriber / Network management, Bayesian Network, Anomaly Indicators.**

## 1. INTRODUCTION

The term anomaly refers to an outlier and to suspicious data that can easily be spotted in small datasets but is hidden and requires intelligent detection in the massive amounts of call data in telecommunications environments. Common inconsistencies in call data are caused by customer churn or attrition, failure in networks [2], potential fraud [3], [4], [5], deliberate or unintended expensive mistakes made by

telecommunications' workers, bad debt risk, or even improper disconnection of communication lines. Malicious behaviour has been noted as one of the key causes for such anomalies, which have consequently led to unquantifiable loss of revenue to many telecommunication networks world-wide.

The most common cause of such malicious behaviour is potential fraud, because it is committed intentionally and it is difficult to combat [3], [4], [5]. To complicate the situation further, the network carriers often do not want to admit that fraud exists as an anomaly in their systems, so that their subscribers do not suspect that fraud is a significant problem. If they do acknowledge it as a significant problem, it might cause churn or cause more subscribers to try to commit fraud. Instead, they prefer this problem to be solved confidentially. If the subscribers suspect fraud but nonetheless keep paying for debts that they did not incur and for services that they did not receive, and if the Telecommunication Service Provider (TSP) does not find a solution to the problem, these customers may decide to seek alternative competing service providers. Furthermore, we have noted that what characterises subscribers' calls is the transformation of their behaviours which appear in a non-linear order (irregular pattern). For these reasons, it is obvious that anomaly prevention security is defeated, which necessitates improvements of detection techniques.

Service Providers are greatly challenged when subscribers, who feel uncomfortable with their services including for example; over-charging, not answering queries promptly, or generally bad customer service, change to competing carriers. It results in the loss of revenue, which is mostly attributed to anomalies. This could quickly put a telecommunication company out of business. Currently, one of the prevailing methods for combating anomalies is block crediting [6], [7]. This means that a bill is sent to the customer who either approves or disapproves the billed amount. This can result in arguments which could frustrate either party. More so, this approach is expensive and may contain errors. Existing approaches to anomaly detection includes for example: rule-based systems [6], statistical systems [4], neural networks [8], Bayesian networks applied to pre-classified anomalies [7], and distance-based systems [9]. These techniques are powerful but their appropriateness can be improved, if every subscriber's anomaly indicators (or significant call attributes) are used. Anomaly indicators

can be identified from probabilistic models for individual subscribers. This is necessary due to the high non-linearity of subscribers' behaviours as a result of explosive technological advancements in phone services. In principle, the Call Detail Records (CDR) that describe subscribers' user profiles generally contain relevant attributes, such as a caller's number, location, duration, destination number, date and time of call. However, most of the methods listed above place equal weight on all these attributes, even though some carry more information than others. This unequal weighting needs to be considered in anomaly detection. If these methods mentioned above make special use of these indicators, the appropriateness of their detections will be improved and the telecommunication revenue losses will be significantly reduced.

Formally, anomaly indicators can be identified from significant call attributes by using the degree of causalities of nodes [10]. It varies from one subscriber to another since they generally have different behaviour. Since there is a massive amount of telecommunication data which requires scalability, and as there are large numbers of subscribers, in practice the existing approaches have many false alarms. Consequently, we propose intelligent agent-based [11] Bayesian networks with unsupervised learning. The positive responses from early users of this methodology, indicates that these novel techniques can be successfully applied to telecommunication markets. David Collings [12] used this methodology to study the rate of adoption of telecommunication services. Other related applications are recorded by [13], [14], and [15]. Our primary objective is to identify significant attributes to facilitate anomaly detection techniques.

In this paper, section one contains the introduction, sections two and three describe the background studies, section four describes the methodology, section five presents our implementation results and interpretations, sections six and seven consist of conclusions, recommendations and future work.

## 2. BAYESIAN NETWORK MODEL

A Bayesian network model is a multivariate probability [16] distribution that is used to define qualitative and quantitative relationships between random variables  $x_1, \dots, x_n$ . The relationships describe the causalities or dependencies, which are represented as a chain rule of joint probability density by expression 1,

$$p(x_1, \dots, x_n) = \prod_{i=1}^n p(x_i | \pi_i) \quad (1)$$

where  $\pi_i$  represents a set of parent(s) of  $x_i$ . Formally, a Bayesian belief network model is a directed acyclic graph expressed as,  $G = \{N(G), E(G)\}$ , where  $N(G) = \{N_1, \dots, N_n\} \Rightarrow$  nodes of graph  $G$  and a set of edges is described as:

$$E(G) \subseteq N(G) \times N(G)$$

The causalities or dependencies of the model  $G$  can be derived from the knowledge of domain experts, obtained from literature descriptions, or mining the structure of the model from data by using unsupervised learning, such as an

evolutionary algorithm. The latter technique is more suitable for studying patterns of call data and the associated anomaly indicators. It is important to know that the most significant node of a model  $G$  contributes crucial information to other nodes. The most significant node can thus be defined as the primary anomaly indicator, which can be identified as the node with the highest number of causalities. In other words, it is a node with the highest number of child nodes (see figures 2 to 5). A child node is at the tip of an edge (or arrow). Moreover, any subsequent nodes with high number of child nodes can also be defined as secondary and tertiary indicators accordingly [1].

## 3. AGENTS AND AGENT ARCHITECTURES

The basic building blocks of the methodology presented in this paper are agent-based models. This term will be further clarified in the subsequent sections. The learning in terms of the Bayesian Network model using a Genetic Algorithm (GA) as discussed below needs to be scalable and distributed through agents.

An agent is anything that can be viewed as perceiving its environment through sensors and acting upon that environment through actuators [11]. In the context of this paper, we use intelligent multi-agent systems, which are component-based software systems. Each of these agents can perceive aspects of its environment and affect that environment either directly or through cooperation with other agents.

Similarly, agents differ markedly from program packages, because of the following characteristics: agents can be distributed, interactional, structured, autonomous or semi-autonomous, and cascaded or run in threads. Thus, when two or more agents share pieces of information, we refer to it as multi-agent systems. How can our evolutionary algorithm be applied in this context?

## 4. EVOLUTIONARY ALGORITHM AND BEHAVIOURAL MODELLING

A Genetic Algorithm (GA) is an evolutionary optimization algorithm that solves problems by eliciting the process of

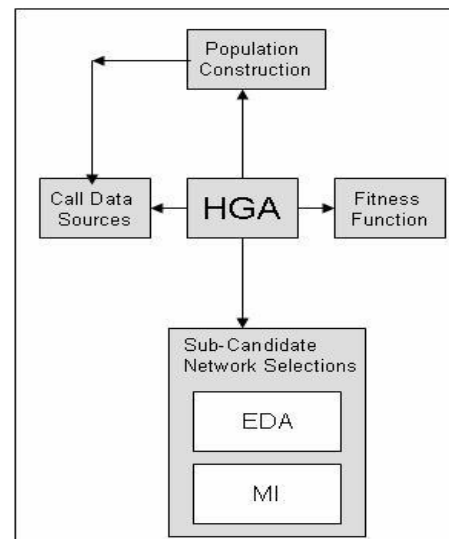


Fig 1. HGA system architecture

evolution [17]. Research has shown that the GA has proved to be one of the best approaches [11], as it is a highly versatile and useful search algorithm to identify relationships and mine structures from data. Similar to other GAs presented in the literature, [18] summarizes a pseudo-code for genetic algorithms as follows:

---

### Pseudo-code for Genetic Algorithms

---

1. **Let  $g = 0$ , be the generation counter**
2. **Initialize a population  $C_g$  of  $N$  individuals;**

$$C_g = \{\vec{C}_{g,n} \mid n = 1, \dots, N\}$$
3. **While no convergence**
  - **Evaluate the fitness of each individual**

$$\vec{C}_{g,n} \in C_g$$
  - **Perform cross-over**
    1. **Select two individuals  $\vec{C}_{g,n1}$  and  $\vec{C}_{g,n2}$**
    2. **Produce offspring from  $\vec{C}_{g,n1}$  and  $\vec{C}_{g,n2}$**
  - **Perform mutation  $\vec{C}_{g,n1}$  and  $\vec{C}_{g,n2}$** 
    1. **Select one individual  $\vec{C}_{g,n}$**
    2. **Mutate  $\vec{C}_{g,n}$**
  - **Select the new generation  $C_{g+1}$**

It is noted that this iterative scheme converges when the candidate network with the best fitness emerges. Crossover is generally considered as the most important genetic operation; this refers to the process of creating a new individual through the combination of genetic materials of two parents. Similarly to mutation, it introduces new genetic material into an existing individual.

Genetic algorithms efficiently solve problems encountered when mining Bayesian network structures from telecommunication call data. In practice, we implemented these processes as multi-agent systems, as illustrated in the system architecture in Fig 1. The Hybrid Genetic Algorithm (HGA) is the central agent that uses the following information theoretic models and mathematical components as required (see sections, 4A-D). Data Sources in figure 1 pre-processes call data before modelling, the agents and results are presented in section 5A. These components iteratively work together with the HGA to evolve individual Behavioural Bayesian Networks (BBN). Examples of individual subscriber BBN generated from our implementation of the HGA are shown in figures 2 to 5.

#### A Population Construction

Population generation is the first step in a GA from the concept of evolution to the selection of candidates. The individuals at the bottom of the population are referred to as parents, who reproduce offspring. Examples of parents are  $\{x1\}$  and  $\{x2\}$  while an example child is  $\{x1, x2\}$  (see table 2). This population of parents is modified via crossover processes, which elicit sexual reproduction and natural adaptation. The iterative construction of offspring results in new generations.

This paper uses a mathematical power-set lattice to generate populations with the call data attributes shown in table 1. The Population Construction agent iteratively interacts with the HGA based on candidate selections from the solution space.

#### B Mutual Information

Mutual Information (MI) [17], [19] is a smart model that measures the sharing of information between two random variables  $X$  and  $Y$ . The unit of measurement is a bit. The model is shown in equation 2:

$$I(X, Y) = \sum_{x,y} p(x, y) \times \log_2 \frac{p(x, y)}{p(x)p(y)} \quad (2)$$

where,  $p(x, y)$  = joint probability distribution of  $x$  and  $y$  while  $p(x)$  and  $p(y)$  are the marginal probabilities of  $x$  and  $y$  respectively. The variables  $X$  and  $Y$  correspond to the call attributes in table 1.

#### C Shannon's Information Theorem

The original Extended Dependency Analysis (EDA) is a reconstructability [20] analysis technique that evaluates the level of significance that exists between variables of a training set. In addition to the original EDA, our methodology includes the mutation process of the genetic algorithm to reproduce a new candidate network; it furthermore, finds the dependency between the two variables using the theorem of Shannon Information Content [21], and produces the winning candidate network. In Fig 1, the EDA receives a sub-set of offspring from the HGA and produces a candidate network. It passes candidates of not more than two variables at a time to the MI. Shannon defines information content,  $h(x)$  as a natural measure of

an outcome  $x$ , in equation 3.

$$h(x) = \log_2 \frac{1}{p(x)} \quad (3)$$

So, for all possible outcomes of  $x$ , the information content is defined in equation 4.

$$h(x) = \sum_x \log_2 \frac{1}{p(x)} \quad (4)$$

It is measured in bits. The smallest number of bits possible to store the information of  $x$  effectively is used. For example, the information content represented using **67.51** bits is preferred to using **98.31** bits to represent the same information.

#### D Fitness Function

One of the important components that make a GA so robust is its fitness function, which measures the optimality of a candidate Bayesian network in our research. The Minimum Description Length (MDL) [16] is suitable to measure the fitness of Bayesian networks to datasets. Since the objective function of our agents is to minimize the number of bits required to store a dataset and its corresponding Bayesian Network, the HGA iteratively invokes the MDL until a minimum score is obtained. In principle, the MDL is made up of two components as shown in equation 5;

$$L(D, B) = \sum_{i=1}^m \log_2 \frac{1}{p(v_i)} + \frac{|B| \log_2 m}{2} \quad (5)$$

where  $D$  = training dataset (call data),

$L$  = length of bits required to store both the dataset and its corresponding network.

$|B|$  = number of parameters in  $B$ , as defined in equation 6;

$$B = \sum_{i=1}^N (j_i - 1) \cdot k_i \quad (6)$$

where  $j_i$  and  $k_i$  are the cardinalities of the child node and its parents set respectively. Also  $N$  = all nodes in a network and  $m$  = number of samples. However,  $\frac{\log_2 m}{2}$  = number of bits that are appropriate to represent each numeric parameter. A candidate network is the network presented to MDL for scoring. Also,  $P(v_i)$  = probability of a sample of  $D$ . This is computed for all values of  $i, \dots, m$ .

That is:  $\mathbf{P(D)} = \mathbf{P(v_1, \dots, v_m)}$ .  $P(v_i)$  is a probability distribution computed using a candidate network.

Therefore, the first component (information content) in expression 7;

$$\sum_{i=1}^m \log_2 \frac{1}{p(v_i)} \quad (7)$$

measures the length (number of bits) required to store the dataset only and the second component in expression 8;

$$\frac{|B| \log_2 m}{2} \quad (8)$$

measures the length (number of bits) required to store the network model. Thus, these two lengths are combined to give the MDL ( $L$ ) of the actual optimal network.

### E Computation of Anomaly Indicators

Having analyzed the Bayesian modelling for subscriber behaviour and using the theories of causalities (see background in section 2), we present a qualitative measure for estimating anomaly indicators (or most significant nodes) on the networks. By definition, the maximum significant node is computed from equation 9.

$$Max(V_k) = \sum_{j=1}^n c_j + [\epsilon], \quad \forall V_k \in G \quad (9)$$

where  $c$  is the number of causalities for every node on the network  $G$ . Also, the  $[\epsilon]$  is optional when two node measures are equal, and it depicts the number of independencies that may be caused by the removal of the node. Some of the related efforts to identify interesting nodes can be seen in [22], but they mostly rely on prior knowledge of domain experts. However, the telecommunications domain expert knowledge may fail due to the inability to update knowledge about the current trends in the telecommunication calls and also due to large datasets.

## 5. EMPIRICAL VALIDATION RESULTS

We believe that existing approaches are not capable of identifying anomaly indicators as pre-requisites for anomaly detection in telecommunications. In our research, we used agent-based behavioural modelling to facilitate anomaly

detection. To validate the universality of our architecture, we have implemented and tested this with real world land-line call data for TSP subscribers. By masking the subscribers' phone numbers, we have ensured the confidentiality of network carriers based on the policies to protect telecommunication customers.

It is necessary to pre-process call data before modelling subscribers' behaviour with the HGA.

### A Pre-processing Model

The following are the cascaded agents that implement this model: the filtering agent, discretization agent and user-profiling agent. The functionalities of the filtering agent were tested using more than 160 subscribers, and more than 73 000 calls with 9 attributes. Subsequently, the output of the filtering agent serves as the input of discretization, and its functionalities maintain discrete values as discrete intervals, and continuous values as continuous intervals. Similarly, the user-profiling agent receives output from the discretization agent and produces an informative call dataset with the attributes and node names for every subscriber (see Table 1).

According to TSPs, the peak period calls are made in the morning, afternoon and evening, while the off-peak period is over-night. We excluded originating-No from modelling a subscriber's behaviour because it does not exhibit any causal relationships to other nodes. The resulting training set is used by the HGA agents to mine individual BBNs.

### B Population Space

The nodes in table 1 are used to generate the population from one generation to another through the crossover processes discussed in section 4A. The following results in table 2 thus formulate a solution space for candidate network models.

Table 1: Call Attributes

Attributes	Node names
originating-No	x0
Call-date	x1
Duration	x2
Destination-no	x3
Destination-net	x4
Location	x5
Call-cost	x6
Peak/off-peak	x7

Table 2: Population Space

{x6,x7}
{x5,x6} {x5,x7}
{x4,x5} {x4,x6} {x4,x7}
{x3,x4} {x3,x5} {x3,x6} {x3,x7}
{x2,x3} {x2,x4} {x2,x5} {x2,x6} {x2,x7}
{x1,x2} {x1,x3} {x1,x4} {x1,x5} {x1,x6} {x1,x7}
{x1} {x2} {x3} {x4} {x5} {x6} {x7}

### C Optimal Bayesian Network Models

A few of the subscribers' models after running the HGA with the call data are shown below. The back-end model results can be presented in XML-BIF, database, CSV (comma separated variable) or text file formats. We used JavaBayes as a front-end to visualise the models.

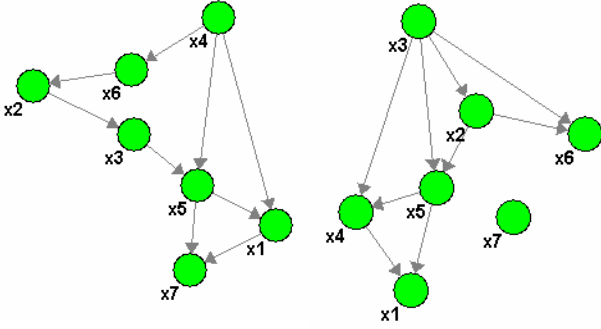


Fig 2: Subscriber-1

Fig 3: Subscriber-2

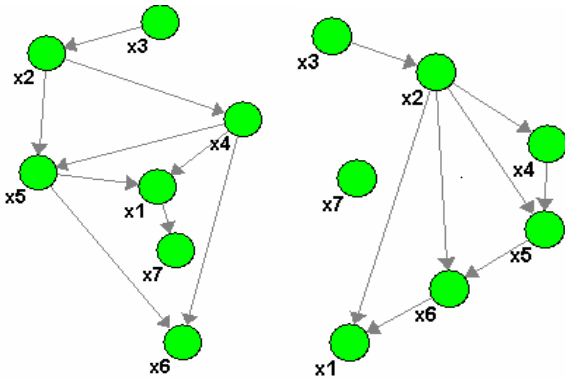


Fig 4: Subscriber-3

Fig 5: Subscriber-4

### D Identification of Anomaly Indicators and Interpretation of Results

With reference to equation 9, the most significant nodes referred to as the most interesting attributes are **x4**, **x3**, **x4** and **x2** in Figures 2, 3, 4 and 5 respectively. Observe that variable **V** in equation 9 represents attributes **x** in figures 2 to 5. The numbers of causalities (see equation 9) for these nodes or attributes on the networks are: **3**, **4**, **3** and **4** respectively. Table 3 shows the significances (causalities) of the call attributes. The values in bold are the anomaly indicators for the corresponding figures.

This means for example, that special focus must be made on **x4** (Destination network) when identifying anomalies for subscriber 1, in figure 2. We can see that, for this individual subscriber, this node carries a lot of information that can be very helpful for detecting anomalies. Moreover, if there is a common most interesting attribute to the majority of the subscribers, it will be important to also consider it together with the usual longest call duration that is normally used in existing approaches.

### E Performance Evaluations

We present the performance evaluations of our implemented hybrid genetic algorithm that finds anomaly indicators (or

most interesting attributes) in telecommunications call datasets. We have conducted a number of experiments to evaluate the performance of our modelling architecture by using publicly available data such as datasets in the UCI [23] repository. Some examples of the evaluation results are shown in table 4.

In table 4, #Attributes refers to the number of attributes present in the dataset. The most interesting attribute is the column that carries a lot of information that is required by other attributes in a Bayesian network model. The most interesting attribute is referred to as an anomaly indicator in telecommunications call data. The #Causalities column is the maximum number of attributes that are depending on the most interesting attribute. The time (s) is the computational time, in seconds, to find the most interesting attribute from a dataset.

Table 3: Significance of Call Attributes

V	Fig. 2 causality	Fig. 3 causality	Fig. 4 causality	Fig. 5 causality
x1	1	0	1	0
x2	1	2	2	<b>4</b>
x3	1	<b>4</b>	1	1
x4	<b>3</b>	1	<b>3</b>	1
x5	2	2	2	1
x6	1	0	0	1
x7	0	0	0	0

Table 4: Performance Evaluation of HGA for finding most interesting attributes from UCI datasets.

Datasets	#Attributes	Most Interesting Attribute	#Causalities	Time(s)
Iris	5	Petallength	4	1
Hardware	7	Class	6	3
Nursery	9	Class	6	5

## 6. CONCLUSION AND RECOMMENDATIONS

The identification of telecommunication anomaly indicators is worth studying as a first step to facilitating anomaly detection. It is clear from our results that these indicators determine many other nodes in every subscriber model. However, having a different model for each subscriber is an important contribution as it reveals variability in call behaviours. It is however, a challenge, which we successfully addressed in this research in order to achieve improved detection quality and decision making.

We therefore strongly recommend that all existing anomaly detection methods should identify and incorporate interesting nodes in their approaches. This will maximize positive detection rates and minimize false alarm rates.

## 7. FUTURE WORK

Our future research direction is to investigate an appropriate learning technique to compute the quantitative knowledge of identifying anomaly indicators using unsupervised algorithms. This can be done by means of probabilities, which can be used to make Bayesian inferences to detect

and predict anomalies in call data. Also, we hope to collaborate with a mobile telecommunication service provider for further validation and evaluation of this novel technique.

#### ACKNOWLEDGEMENT

We are grateful to the entire management and staff of Complex Adaptive Systems (Pty.) Ltd. and the African Institute for Mathematical Sciences (AIMS) South Africa, for funding this research to a successful completion.

#### REFERENCES

- [1] P. Burge, J. Shawe-Taylor, Y. Moreau, B. Preneel and C. Stoermann. "Fraud Detection and Management in mobile telecommunications networks", In *proc. of the European Conference on Security and Detection*, 1997.
- [2] K. Sequeira and M. Zaki. "Anomaly-based data mining for intrusions". In *ACM SIGKDD 02*, 2002.
- [3] S. Rosset, U. Murad, E. Newmann, Y. Idan and G. Pinkas, "Discovery of fraud rules for telecommunications challenges and solutions", In *Proceedings of the 5th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 409--413. ACM Press, 1999.
- [4] H. Michael, D. Lambert, C. Pinheiro and D. Sun, "Detecting Fraud in the Real World", Technical Report, Lucent Technologies, 2000.
- [5] I. Osunmakinde and A. Potgieter. "Telecommunications Fraud Detection using Bayesian networks". An essay submitted to the African Institute for Mathematical Sciences (AIMS), South Africa, 2005.
- [6] T. Fawcett and F. Provost, "Adaptive Fraud Detection" *Data Mining and Knowledge Discovery* 1, 2, Kluwer Academic Publishers, Boston, pp. 1-28, 1997.
- [7] J. Hollmen and V. Tresp, "Call-based fraud detection in mobile communication networks using a hierarchical regime-switching model". *Advances in Neural Information Processing Systems, 11<sup>th</sup> proceedings of the 1998 Conference (NIPS' 11)*, MIT press, pp 889-895, 1999.
- [8] C. Bounsaythip, E. Rinta-Runsala, "Overview of Data Mining for Customer Behaviour Modelling", Technical Report, TTE1-2001-18, VTT Information Technology, Finland, 2001.
- [9] D. Bay and M. Schwabacher, "Mining distance-based outliers in near linear time with randomization and a simple pruning rule". In *proc. of 9<sup>th</sup> annual ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2003.
- [10] J. Pearl, "Probabilistic reasoning in Intelligent systems:", *Networks of Plausible Inference*, Morgan Kaufmann Publishers, 1988.
- [11] S. Russel and P. Norvig, "Artificial Intelligence, a modern approach", 2<sup>nd</sup> edition, Prentice Hall, New Jersey, 2003.
- [12] D. Collings, "Agent based customer modelling", *Computing in Economics and Finance*, Society for Computational Economics, No 1352, 1999.
- [13] B. LeBarron, "Building financial markets with artificial Agents", *Computational Markets*, MIT Press, Cambridge MA, 1999.
- [14] M. Janssen, "An integrated approach to simulating behavioural processes", *Journal of Artificial Societies and Social Simulation*, Vol. 2 No. 2. 1999.
- [15] D. Bunn and F. Oliveria, "An application of agent-based simulation to the new electricity trading arrangements of England Wales", *IEEE Transactions on Evolutionary Computation*, 1999.
- [16] N. Nilsson, "Artificial Intelligence: ", A new synthesis, first edition, San Francisco USA. Morgan Kaufmann Publishers, 1998.
- [17] M. Gell-Mann, "The Quark and the Jaguar, Adventures in the Simple and the Complex", Abacus, Lancaster, London, 2001.
- [18] P. Engelbrecht, "Computational Intelligence", John Wiley, England, 2002.
- [19] A. Papoulis, "Probability, Random Variables, and Stochastic Processes", 2<sup>nd</sup> edition, New York, McGraw-Hill, 1984.
- [20] M. Zwick, "Wholes and parts in General Systems Methodology", *The Character Concept in Evolutionary Biology*, Academic Press, 2001.
- [21] D. Mackay, "Information Theory, Inference and Learning Algorithms", Cambridge Press, 2003.
- [22] S. Jaroszewick, "Interestingness of Frequent Itemsets Using Bayesian Networks as Background knowledge", *proc. of the 10<sup>th</sup> ACM SIGKDD international conference on knowledge discovery and data mining, ACM press*, 2004.
- [23] UCI MLRCS., UCI Machine Learning Repository Content Summary. <http://www.ics.uci.edu/~mllearn/MLSummary.html>.

**Isaac Olusegun Osunmakinde** is completing his Masters in the Agents Research Group (ARG) of the Computer Science Department, University of Cape Town, South Africa. He has a post graduate diploma (pgDS) in Mathematical Sciences from the University of Stellenbosch, South Africa, a B.Sc. (Honours) in Computer Science from UNAAB, Nigeria, a HND in Computer Science and a ND in Computer Technology from YABATECH, Nigeria. He is a member of the IEEE Computational Intelligent Society and an Oracle certified professional (OCP). His current research interests in AI are probabilistic and mathematical modelling.

**Anet Potgieter** is a senior lecturer and head of the Agents Research Group of the Computer Science department, University of Cape Town, South Africa. She holds a Ph. D. (Computer Science) from the University of Pretoria (South Africa). Her current research interests include complex adaptive systems, distributed artificial intelligence, sensor networks and software engineering.