Neural Network Primer

Issue 2.0 April 2000
Cerebrus is the direct successor to SuperSleuth™, Nortel Networks Fraud Solutions advanced fraud management offering. The change in name has been necessary to secure an identity that may be used consistently throughout the global fraud management market.

Cerebrus 3.1, the first release under the new identity, provides continuity of SuperSleuth 3.0’s market leading capabilities and benefits because only its identity has changed. Additionally it delivers new features that make the management of pre-paid fraud more effective and which enhance the Hot List and CDR rating features.

About Nortel Networks Fraud Solutions

This Neural Network Primer is prepared by Nortel Networks Fraud Solutions.

Fraud Solutions is a business established by Nortel to offer telecommunications operators fast, effective and dependable solutions to enable them to control and reduce the costs of fraud.

Nortel is an acknowledged world leader in the field of public and corporate telecommunications solutions, well known and respected for its innovation and quality. The creation of Fraud Solutions is the natural extension of Nortel’s approach to meeting the global demands of the telecommunications industry. Fraud Solutions’ mission is to apply advanced technologies to allow telecommunications operators to protect their business interests and profitability threatened by fraud.

Fraud Solutions provides operators with a focused, dedicated team of specialists able to provide practical and proven solutions to the problem of fraud based on the range of Cerebrus software products.
Table of Contents

1. INTRODUCTION 4
   1.1. HISTORICAL BACKGROUND 4
   1.2. WHAT IS A NEURAL NETWORK? 4

2. HOW DOES AN ARTIFICIAL NEURAL NETWORK WORK? 5
   2.1. TRAINED NOT PROGRAMMED 5
   2.2. THE ANATOMY OF AN ARTIFICIAL NEURAL NETWORK 5
   2.3. TYPES OF ARTIFICIAL NEURAL NETWORK 8
   2.4. ARTIFICIAL NEURAL NETWORK TRAINING 8
   2.5. HOW DOES AN ARTIFICIAL NEURAL NETWORK REACH A DECISION? 8

3. WHEN SHOULD AN ARTIFICIAL NEURAL NETWORK BE USED? 10
   3.1. PROCESSING COMPLEX, RULE-DEFYING PROBLEMS 10
   3.2. ADAPTING TO CHANGING CIRCUMSTANCES 10
   3.3. AN ARTIFICIAL NEURAL NETWORK SYSTEM CAN EXPLAIN ITS RESULTS 10

4. WHAT IS CEREBRUS? 11
   4.1. CEREBRUS DETECTS TELECOMS FRAUD 11
   4.2. NEURAL NETWORKS BASED SUBSCRIBER PROFILER 11
   4.3. CONVENTIONAL FIRST LINE DEFENCES 11
   4.4. CALL DETAIL QUERIES DATABASE 12
   4.5. ROAD MAP 12

5. FURTHER INFORMATION ON NEURAL NETWORKS 13
1. **Introduction**

1.1. **Historical background**

First implemented in the early 1960s neural networks only began to develop significantly in the mid 1980s with the introduction of new neural network architectures and advances in processing technologies. Since then, they have been successfully used in a variety of industries such as Finance, Retail, Manufacturing, Energy, Health, Telecommunications and Security. Their application ranges across many fields such as financial market prediction, sales forecasting, mineral exploration, process control, speech recognition, marketing and, of course, fraud detection.

1.2. **What is a neural network?**

Artificial neural networks are a form of computation that is modelled on brain processes.

In the brain, billions of neurons are interconnected through synapses to form a biological neural network. Information is transmitted through the network by electrochemical signals. In artificial neural networks this process is simulated.

There are many types of artificial neural networks (or ANN) but they all share common characteristics.

- They consist of connected units that represent neurons. Each neuron receives inputs, processes them and generates an output.
- The units are arranged in specific architectures with specific patterns of connections between the units. These represent the synapses.
- The network is trained to respond to data presented rather than programmed.
2. How does an artificial neural network work?

2.1. Trained not Programmed

The principal difference between neural networks and standard computation methods is that the neural network learns to perform a task through training, whereas in standard computation this is achieved by programming.

Many real-world situations are difficult to model with rules and algorithms that are essential to the success of conventional programming techniques. This is partly because of their complexity, but also because it is impossible to specify exact rules. This is particularly so in situations involving human judgement. Decisions are taken on the balance of evidence. The importance of one factor may depend on many other factors. Human judgement in such circumstances is the result of experience. Neural networks imitate this process by learning to recognise patterns of behaviour.

Rather than having to define rules and algorithms to describe a problem and then writing a specific program for that application, neural networks allow a quite different approach. During the training process, they learn to generate the correct response. In this process example data is presented and a feedback loop applies correction or affirmation, modifying the neural network’s response to reinforce the desired behaviour.

A neural network is a general tool that is effectively characterised by its training. An analogy is a person learning to speak a foreign language. People have a general capability to learn a language. As children, we learnt to speak our mother tongue primarily by example and yet we became fluent. Thus we became English speaking, French speaking etc. Some have learned additional languages, by training as opposed to the re-programming of their brains. These people can be characterised as bi-lingual or multi-lingual and the language spoken.

2.2. The anatomy of an artificial neural network

The operation of the various neural network architectures is quite different but all share some features.

All networks consist of two or more layers of units (neurons). These are linked by weighted connections. Figure 1 provides a simplified model of an individual neuron, in cartoon form. The input signals here are received and are weighted according to the value assigned to each input.

The neuron is triggered by its inputs where, in the simplified model, a value of greater than five (5) has to have landed in the hopper. When the neuron is excited, it generates its output signals. These are sent to the next layer of neurons. In reality, an artificial neuron (unit) would transmit both negative and positive signals dependent upon the value and weighting of its inputs and the internal algorithm, represented here by the hopper mechanism. The actual algorithms used are arithmetic in nature and are more complex than the simple threshold suggested in the cartoon.
Figure 2 demonstrates that every parameter in the input data set (i1, i2 & i3) received by the input layer is presented to every neuron unit in the hidden or middle layer of the network. The stars represent the transmission of the signals, for clarity only some of the signal paths are de-emphasised, but all paths are active. Thus, each unit's output is dependent upon all of its inputs.
The hidden layer units each respond to the incoming signals according to the weight of its input (represented in the diagram by the line thickness). The signals are modified by the weights of the connections, which will tend to strengthen or weaken signals that pass along the connection. Each hidden layer unit then transmits a signal to each unit in the output layer (Figure 3).

This transmission of signals through the network eventually results in particular values being output by the final layer of output, shown in Figure 4, where the result is represented by stars.

---

**Figure 3** The input layer has been activated and its signals are propagated

**Figure 4** The neural network's decision is presented
If the network has been architected and trained to identify fraudulent behaviour compared to non-fraudulent behaviour, as is the case for Nortel Networks Fraud Solutions's Cerebrus for instance, then there will probably be two output units. One of these will register cases of fraud, the other non-fraud. During training, the connection weights are adjusted so that fraud cases will result in a large value output by the fraud node and a low value at the non-fraud node.

2.3. Types of artificial neural network

There are several different neural network architectures but all are either supervised or unsupervised. This refers to the training method used.

The supervised systems, such as Cerebrus, are trained to respond to certain inputs by using training data of known behaviour. Once the system has trained it is able to generalise and identify inputs by their resemblance to its training set. The training process requires that the result for each example input data is confirmed or corrected until the desired level of accuracy is achieved. Most of the commercially successful applications use supervised learning where it is possible to incorporate known behaviour into the network's learning and to refine the network performance during the training process.

Unsupervised neural networks are able to automatically classify input data as belonging to a set of similar input data. Unlike supervised networks, training does not involve the supervised presentation of examples and correction of results. The system automatically assesses the characteristics of the data itself and it is necessary for the system user to analyse its results in order to understand and describe the identified behaviour. They are powerful tools with which to identify and classify behaviour where the nature of the information being sought is otherwise unknown.

2.4. Artificial neural network training

Cerebrus uses supervised training where the training data consists of many examples of both fraud and non-fraud cases. This data is presented to the network and the resultant output compared to the desired output. The desired output being fraud when the input data is an example of fraud, for instance. The weights in each neuron unit are automatically adjusted at each stage to bring the actual output more into line with the desired output. This process of continual adjustments over repeated presentations of the training inputs eventually results in a trained network.

The Cerebrus network uses a combination of advanced techniques to speed up the learning process. It is also unique in that this process is automated to enable the neural network to autonomously incorporate operational results and so adapt to changing behaviours over time. In this way, optimum performance can be automatically maintained without recourse to periodic re-training by a human expert, as will typically be the case for conventional neural networks.

2.5. How does an artificial neural network reach a decision?

A graphic illustration of the decision process can be given using the example of handwriting recognition. In a supervised network of the type used in Cerebrus, many examples of handwritten characters are presented to the network in training. If there are 26 characters to identify then the network will usually have 26 output units. The neural network will be trained
so that each output unit will respond to a particular character. The weights on every input in
the network will be have been set by training such that the appropriate output unit will
present a large value at its output, while the other 25 will give lower value outputs.

When handwriting is presented to the trained network, the output unit with the largest signal
identifies the character being presented. In the example shown in Figure 5, parameters
which represent the hand written character "b" are presented to the input layer. The input
units send signals to the hidden layer units, which in turn send signals to the output layer
units. The strength of signals that are sent by each hidden layer unit are dependent upon the
weights established during training. Each output layer unit receives a signal from every
hidden layer unit, and the strength of the signal it outputs is dependent upon its weights. In
the example, the weights are set so that the "b" output unit delivers the largest signal,
indicating that the hand-written character is "b".

In some cases, patterns will give intermediate values. These will indicate that the network
cannot classify the pattern precisely. In practice, dictionaries and grammars to deal with
ambiguous cases would support the system.

Figure 5 Character Recognition
3. When should an artificial neural network be used?

3.1. Processing complex, rule-defying problems

Neural networks often provide the best solutions in situations where it is difficult to establish hard and fast rules, and the data to be analysed is complex. The more complex the data the greater the advantage of the neural network. This is shown particularly in their use for computer vision systems where the input is exceptionally complex. Once properly trained and established neural networks can give very accurate results.

Because of their arithmetic nature, neural networks are also good at processing large volumes of data. Computers are essentially designed to deal with arithmetic operations and so neural networks are computationally very efficient, much more so than procedural operations. Also neural networks identify patterns in data rather than systematically analyse it, which also lightens the processing burden as compared to procedural, rule based analysis methods.

3.2. Adapting to changing circumstances

Neural networks are particularly suitable for adaptive systems; systems that learn with experience. As more data is accumulated, the network performance can be monitored and if necessary, the network can be retrained using the new information. The ability to respond to changes of behaviour over time and the ability to quickly process large volumes of data make fraud detection an ideal application for neural networks. Here the behaviour of both the fraudster and the legitimate customer are always changing.

By comparison, procedurally based systems have difficulty in responding to change. When circumstances have changed beyond their tolerance re-analysis and re-implementation is required.

3.3. An artificial neural network system can explain its results

It is often claimed that neural networks have limited use, because it is difficult to understand why they have given a particular result. The reason given for this difficulty is that their decision is based upon the relative influences of the weightings used in each neuron unit. It is thought that because these are embedded within the network, and are not externally visible, it is not possible to understand the network's outputs.

While this may be true for unsupervised neural networks that identify and classify previously unknown patterns, it is not necessarily the case for supervised networks.

Supervised networks are targeted at specific tasks, such as identifying if a telephone subscriber's behaviour is likely to be fraudulent or not. The characteristics of the basic problem being solved are known. The reason for applying a neural network is the rule-defying complexity and subtle variations in behaviour profiles, as well as the volume of data involved. The task is to reliably identify a few problematic subscribers amongst the millions of good subscribers.

Under such circumstances, where a supervised neural network is tasked with identifying specific conditions, it is possible to indicate what has caused the neural network to give a particular result. Providing a summary of essential input parameters allows a human operator to rapidly and easily understand why a particular data set has led to the specific result.
4. **What is Cerebrus?**

4.1. **Cerebrus detects telecoms fraud**

Cerebrus is a multi-technology, or hybrid, system targeted at the identification of telephone fraud. At Cerebrus’s heart is the ability to rapidly detect telephone fraud by monitoring subscriber behaviour through the advanced application of neural network technology.

4.2. **Neural Networks based Subscriber Profiler**

Cerebrus creates individual behaviour profiles for each individual subscriber. Mimicking the analyst, Cerebrus monitors these profiles and detects anomalous activity patterns that suggest fraud. The operator is then able to further investigate the identified high risk subscribers using the information provided by the Profiler’s “Reasons” display, First Line Defences and Call Information Queries modules. The analyst provides feedback so that the system can continually and automatically increase its knowledge. Because each individual profile is unique to each subscriber, Cerebrus is able to track subscriber behavioural changes, allowing legitimate evolution without analyst intervention. This is key as behaviour changes and fraudulent behaviour for one subscriber may look like normal behaviour for another. Additionally the neural network is able to identify and react to new types of fraud that it has never seen before. The Profiler may also be expected to identify frauds that have been adapted not to trigger thresholds.

New customers have no history and as such represent a high risk. Cerebrus identifies new customers as they begin to use the network. In many cases Cerebrus will detect new customer activity before the analysts are informed of their existence through normal channels.

Subscriber Profiler maintains its own detection performance and accuracy by training the neural network whilst the system is on-line. This allows Cerebrus’s Subscriber Profiler to maintain optimum performance 24 hours a day.

The adaptive nature of Cerebrus’s profile based detection can be up to twenty times more accurate than purely conventional systems. It may be deployed in conjunction with the other Cerebrus options or used alone as an enhancement to an operator’s detection and analysis processes.

4.3. **Conventional First Line Defences**

In combination with the Subscriber Profiler the First Line Defences enhance the power and efficiency of Cerebrus. They provide a range of conventional tools that detect call based fraud “tell-tales” reported by subscriber. Thus the analyst is able to identify the fraud “tell-tales” for the high-risk subscribers identified by Subscriber Profiler. This process significantly reduces analyst overload by generating a subscriber-oriented case. The analysts’ effectiveness is enhanced as they concentrate on suspect subscribers rather than huge numbers of false alarms. First Line Defence features include:

- Hot Lists, which detect the use of, suspect originating and destination telephone numbers and codes associated with previous and potential frauds.
Call Collision detection that identifies overlapping calls, a frequent characteristic of Call Selling, Premium Rate Service revenue inflation frauds and surfing frauds such card cloning.

Call duration tests that monitor individual and aggregate calls for specific conditions such as calls being unusually long or short. These can be indicators of calls being involved in revenue frauds or attempts to hack security mechanisms for services such as voice mail and calling cards.

Call length pattern monitoring that detects unusual distributions of call lengths as might be found in a Call Sell operation or through the application of auto-diallers in a PRS fraud.

Call value tests based on nominal tariffs. These detect specific conditions associated with the values of calls employing different services over a range of periods. This can also be used to identify customers who are generating high bills over a short period, a signature of genuine bad debt.

Feature use tests that monitor the use of “fraud friendly” service features such as explicit call transfer.

Independent detection treatments for identified customer groupings.

4.4. Call Detail Queries database

The queries application supplements the Subscriber Profiler and First Line Defences. Based on the Oracle commercial database package it enables the analyst to inspect stored call record information when confirming fraud and developing Hot Lists.

The range of query parameters include:

- Specific and partial originating numbers and codes
- Specific and partial destination numbers and codes
- Call destination and service types
- Call feature use
- Short and long call durations
- Call date and time ranges

4.5. Road Map

Future releases of Cerebrus will introduce

- Enhanced reporting
- Advanced user interface design
- Case management
- Signalling System 7 based fraud detection
5. Further information on neural networks

A very good, readable, discussion of neural networks and their applications aimed at the general reader is:


Two general technical introductions are:


These, like all explanatory texts on neural network theory, require a high level of mathematics to understand them.

There are many good web sites devoted to neural networks. Two good academic sites are:

- *Neural Networks Working Group at*
  http://zsw.e-technik.uni-stuttgart.de/Fachgebiete/SYS/NeuroNetz/nn_mainE.htm
- *Neural Network Resources at*
  http://www.fdc.co.uk/idev/neural.html
More Information

For more information on telecommunications fraud and counter-measures please contact Nortel Networks Fraud Solutions via the Cerebrus Information Line

Europe

Nortel Networks Fraud Solutions
Nortel Networks Corporation
Great Eastern House
Edinburgh Way
Harlow
Essex CM20 2NQ
United Kingdom

Cerebrus Information Line    UK :   +44 (0) 208 945 55 99
                             Fax :   +44 (0) 208 945 54 60

The Americas

Nortel Networks Fraud Solutions
(Mail Stop A0311M)
Nortel CALA Inc.
1500 Concord Terrace
Sunrise
FL 33323-2815
USA

Cerebrus Information Line    USA :   +1 954 858 7493

Asia

Nortel Networks Fraud Solutions
Nortel Networks Malaysia SDN BHD
2nd Floor, Kompleks Penchala
50 Jalan Penchala
46050 Petaling Jaya
Selangor Darul Ehsan
Malaysia

Cerebrus Information Line    Malaysia :   +603 778 62132
                               Fax :   +603 778 62100

E-mail :   ssinfo@nortelnetworks.com
Web Site :   http://www.fraud-solutions.com

All figures, data and other material contained in this publication are typical and must be specifically confirmed in writing by Nortel Networks Corporation before they become applicable to any tender, order or contract.

Nortel Networks Corporation takes every precaution to ensure that all information contained in this publication is factually correct but accepts no liability for any error or omission. No freedom to use patents or other property rights is implied by this publication.

© Nortel Networks Corporation 2000