BitDefender Antispam NeuNet

Whitepaper

Cosoi Alexandru Catalin
Researcher BitDefender AntiSpam Laboratory

bitdefender
secure your every bit

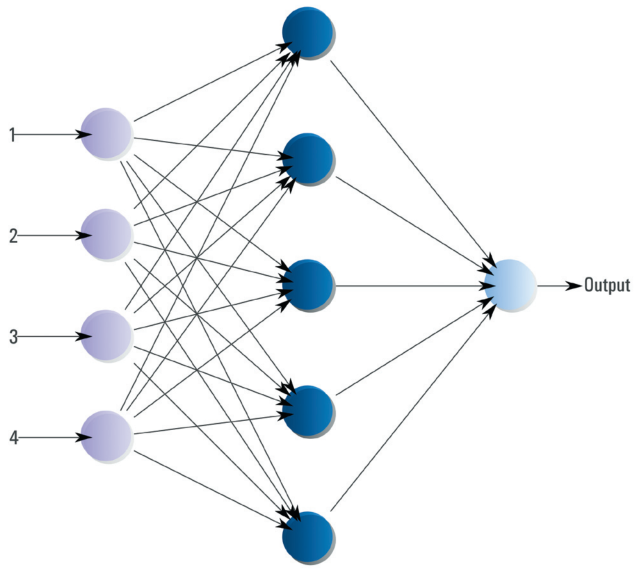
Contents

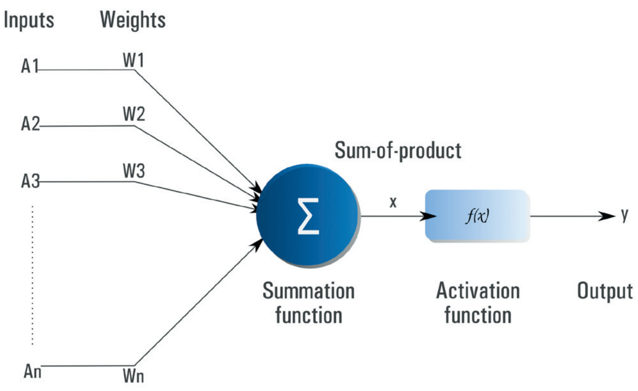
Overview of the Spam Issue

In recent years, electronic mail users around the world have noticed that an increasing amount of unsolicited e-mail reaches their mailboxes. So far, a number of filtering methods have been proposed to address this problem, such as: Bayesian, Black-Lists/White-Lists, Image, URL filtering, Heuristic, etc. The idea behind any spam filtering technique, (heuristic, probabilistic or keyword based) is the same: as spam messages usually look different from legitimate messages, a good way to identify and stop them is to detect these differences. Judging by the results of these filtering methods and given that spam changes every day, the best way to solve this problem would be to use all of these features for a combined and more accurate effect. Easier said than done!. Ever since the emergence of these technologies spammers have improved their techniques, so that spam still gets to its destination. They have used obfuscation, masking words so only a human would understand them, they have taken advantage of the vulnerabilities of Html-parsers or even masked the content in such a way that it would be almost impossible for a computer to make the difference. Anti-Spam solutions have had to increase the frequency of the updates and also to develop more heuristics in less time. The need for an automatic process that would quickly learn the characteristics of the new spam without affecting the accuracy of detection on less recent spam has become vital. The answer we have found to address this problem lies in the artificial neural networks.

About Neural Networks

A neural network consists of a large number of processing elements, called neurons. Each neuron has an internal state, called its activation or activity level which is a function of the inputs it has received. Typically, a neuron sends its activation as a signal to several other neurons. A neuron can send only one signal at a time, although such signal may reach several other neurons. An artificial neural network can also be seen as an information processing paradigm that is inspired by the way biological nervous systems, such as the brain, process information. Neural networks take a different approach to problem solving than that of conventional computers. Conventional computers use an algorithmic approach -i.e. the computer follows a set of instructions in order to solve a problem. Unless the specific steps that the computer needs to follow are known, the computer cannot solve the problem, which restricts the problem solving capability of conventional computers to problems that we already understand and know how to solve. In addition to that, conventional computers use a cognitive approach to problem solving; the way the problem is to be solved must be known and stated in small unambiguous instructions. These instructions are then converted to a high level language program and then into machine code that the computer can understand. These machines are totally predictable; if anything goes wrong it is due to a software or hardware fault. Neural networks and conventional algorithmic computers do not compete but complement each other. Certain tasks, such as arithmetic operations, are more suited to an algorithmic approach whereas others require neural networks. Even more tasks ask for systems that use a combination of the two approaches (normally a conventional computer is used to supervise the neural network) for a maximum of efficiency.



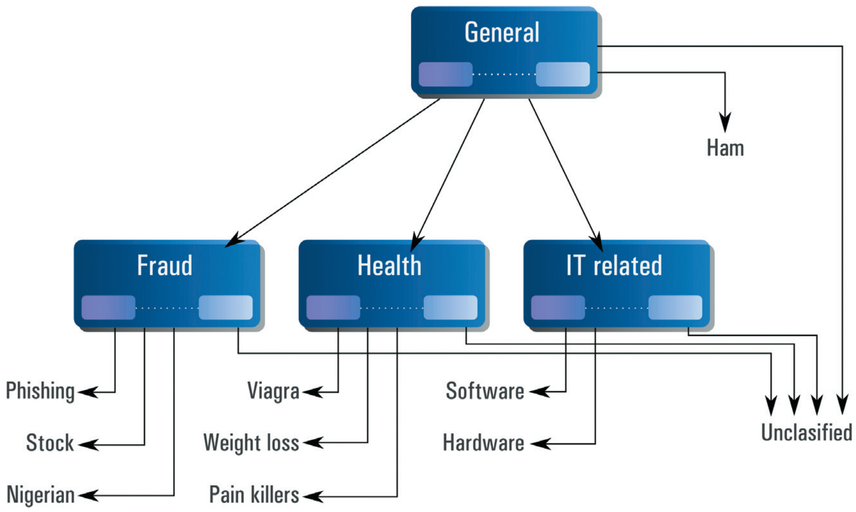


A neural network is said to learn off-line if the learning phase and the operation phase are distinct. A neural network is said to learn on-line if it learns and operates at the same time. Usually, supervised learning is performed off-line, whereas unsupervised learning is performed on-line.

New Structure Using Neural Networks

Our idea was to create an automated process that will gather the spam and ham corpus over a certain period of time, study its characteristics and learn without any human involvement. The faster this process is done, the faster the response.
But neural networks do have some problems. When met with large amounts of data, results tend to drop. Feed-forward neural networks tend to forget some of the information learned at the beginning of the process, or the output data becomes somewhat chaotic. Based on this insight, as well as on the fact that spam can be split into several distinct categories, we have developed a tree of neural networks that can classify large amounts of data faster and without affecting detection accuracy. Each of the neural networks in this hierarchy works on a different type of spam, so that the input and the output are limited enough not to confuse the net and to keep performance at its highest.

A good example of this would be the fact that we have created a subcategory called fraud, which contains messages trying to trick the user into sending money or giving away information about his credit card.
Further, we have identified a subtype of fraud, namely phishing, with its resourceful variants: once, a spammer trying to steal information about the user's credit card by disguising himself as a self claimed rich Nigerian who wants to get some money out of the country and needs your help to do so, by sending him some money; other times it is the lottery trick, where the spammer tries to get personal information from you by claiming that you have won a few million dollars at the lottery; there's also the stock story, where you receive tips on how to buy stock, and so on. It is obvious that all these have something in common and that they can make up a category of their own.
Each subtype has its own characteristics which set it apart, which allows for further subcategorization.



1. Run requested heuristics
2. Extract input vector
3. Reduce noise
4. Pass to the neural network
5. Classify

During the training, if the neural network module does not find a category into which to integrate a particular pattern, it will create a whole new category. Therefore, if too wide a variety of spam is fed to the neural network, the number of categories will increase, which is likely to slow down the analysis. However, if the neural network is specialized on a single type of spam, even with an increased number of heuristics the rapid overflow of output categories is avoided and the analysis is more precise and refined.

Supposing now that the hierarchy of neural networks has been trained and that it is ready to be tested, when an email arrives the system must provide an answer as to its nature: legitimate, spam (a certain type) or “don’t know” (which will also be considered legitimate to prevent false positives). First, the general type of heuristics will be run on the mail to see what category can accept it. If no matching category is found, the mail will be considered legitimate; otherwise, it will be passed on to the next neural network that deals with that type of spam and the algorithm will repeat. At this point, information is extracted, but only as much as needed for the input required for the neural network in question. If the next level the mail reaches is another neural network, the information is passed on and the algorithm repeats. If the mail cannot be classified, it counts as legitimate mail but if the next level is a final category (a leaf in our tree) then that email has been classified and the process ends.

Therefore, the process works based on a selective information extraction basis, which gives a speed boost to the analysis. Moreover, the neural networks approach is more refined and potentially far more accurate and reliable in accomplishing this task.

Efficiency

The detection rate improves consistently when new inputs are added, and can easily grow to almost (or even) 100%. Also, the number of heuristics that can be added is infinite, without being afraid of the time performance. The key element of the detection is not how many heuristics we have, but the patterns that are discovered during the training phase. A certain keyword in an email doesn’t mean that it’s certainly spam, but also doesn’t mean that it isn’t. A pattern will be an entire list of key elements found in the body of the email, and the analysis process can be done even if there is only one word that can be considered spam-like. If during training, a similar email entered the neural net, the analysis process would identify it correctly.

Our experiments show that the neural networks approach is more refined, more mathematical and potentially far more accurate and reliable in accomplishing this task. Using only this filter (BitDefender NeuNet - patent pending technology), on a set of more than two millions of emails (which 80% were used only in training and 20% in testing) we achieved 100% detection on the training corpus and 97.56 on the test corpus, and the system performed much more faster than a heuristic filter. In conclusion we consider this filter as being the next level in fighting spam using neural networks.

About BitDefender®

BitDefender is a leading global provider of security solutions that satisfy the protection requirements of today's computing environment. The company offers one of the industry's fastest and most effective lines of security software, setting new standards for threat prevention, timely detection and mitigation. BitDefender delivers products and services to over 41 million home and corporate users in more than 100 countries. BitDefender has offices in the United States, the United Kingdom, Germany, Spain and Romania. Further information about BitDefender can be obtained by visiting: *www.bitdefender.com*

Contact Info

Efficient communication is the key to a successful business. For the past 10 years SOFTWIN has established an indisputable reputation in exceeding the expectations of clients and partners, by constantly striving for better communications. Please do not hesitate to contact us regarding any issues or questions you might have.

Country: **U.S.A**
Contact: Eric D Lewis
Function: General Manager
Company: **BitDefender LLC**
Address: 6301 NW 5th Way, Suite 3500
Fort Lauderdale, Florida 33309
Phone: 954 776 62 62
Fax: 954 776 64 62
Email: *sales@bitdefender.us*

Country: **Romania**
Contact: Oliviu Talianu
Function: Country Manager
Company: **SOFTWIN SRL**
Address: 5th Fabrica de Glucoza St. Bucharest
Phone: +40 21 2330780
Fax: +40 21 2330763
Email: *sales@bitdefender.ro*

Country: **Germany**
Contact: Martin Siemens
Function: Geschäftsführer
Company: **Softwin GmbH**
Address: Karlsdorfer Straße 56 88069 Tettnang
Phone: 07542/94 44 44
Fax: 07542/94 44 99
Email: *msiemens@bitdefender.de*

Country: **Spain**
Contact: Florin Baras
Function: General Manager
Company: **Constelación Negocial, S.L**
Address: C/ Balmes 195, 2ª planta, 08006 Barcelona, España
Phone: +34 932189615
Fax: +34 932179128
Email: *fbaras@bitdefender-es.com*