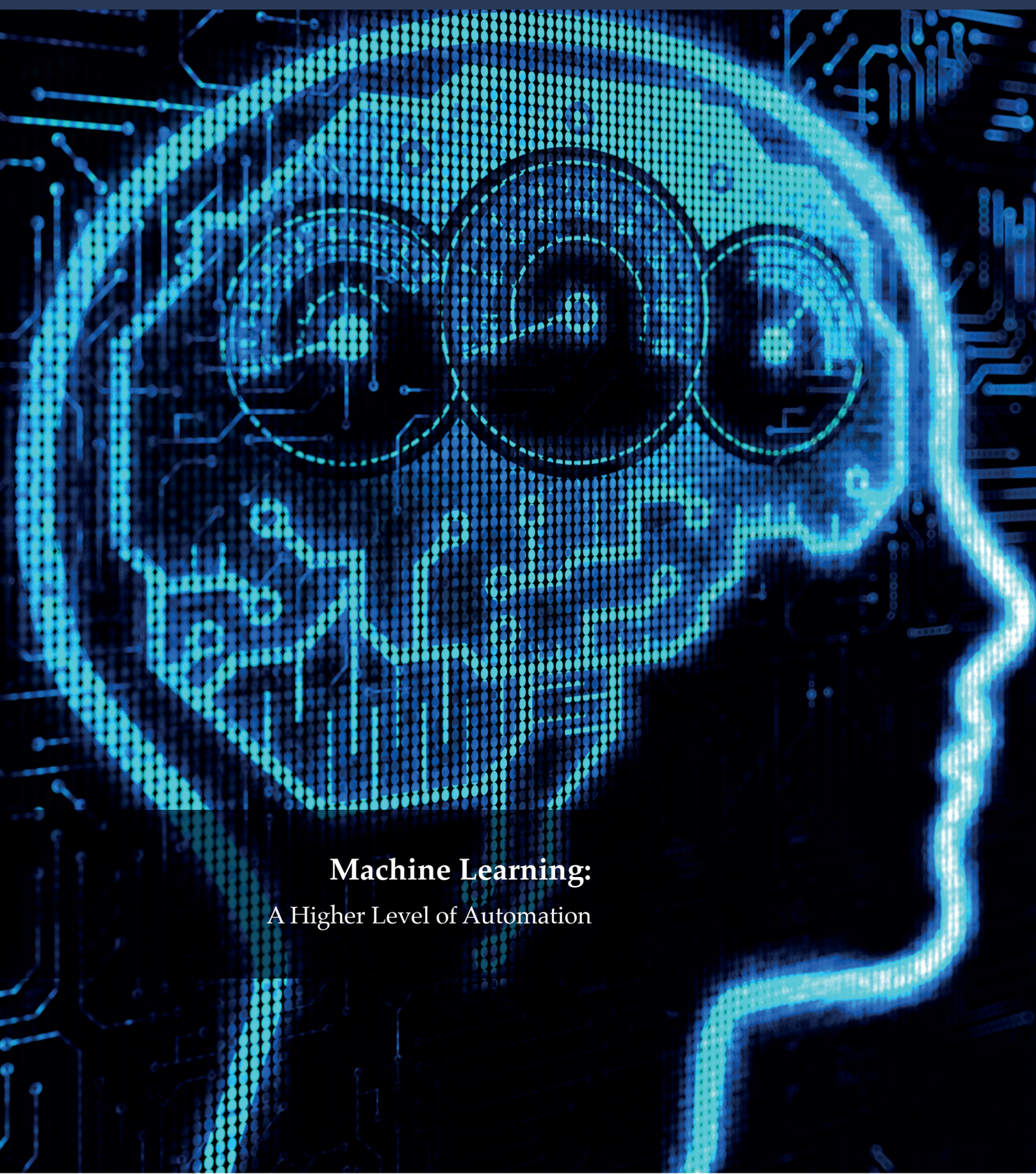
DARKTRACE

WHITE PAPER



Machine Learning:

A Higher Level of Automation

Machine Learning: A Higher Level of Automation

Overview

A new era in cyber has begun. Today, machines are fighting machines, and sophisticated attackers and criminal groups are ready to pounce at any opportunity. The battlefield is a corporate network; the prize is control of the company.

The danger today is not just the classic scenario of information being stolen, or a website being defaced, but the quiet and unseen attack – attackers that creep in and can change your systems at will, or install kill switches, ready to be activated. These attacks are sophisticated, using previously unseen custom code, only crossing the boundary defenses once, never sending information out. They may only be active for a few seconds a year, but when commanded to act, they are fatal.

Legacy approach

Legacy tools are failing to deal with this new threat reality, and many now face extinction. This is because the traditional approach relies on being able to pre-define the threat in advance, by writing rules or producing signatures. In today's environment, this approach is fundamentally flawed:

- Threats are constantly evolving – they do not fit your model or signatures
- Rules, policies and signatures are continually out-of-date – they miss subtle emerging threats
- Insider threat is growing – it is difficult to spot insiders as they are a legitimate presence on the network

The reality is that advanced threats bypass legacy defense tools. New black hat machine intelligences only enter the organization once – from that point of entry, they listen, learn how to behave, how to blend in, and how to appear as authentic as the real devices, servers and users. These automated attackers can hide their actions quietly among everyday tasks. No more brute force port-scanning or head-on attacks. Instead, tiny actions are buried in amongst the noise of normal operations.

New machine learning

Thanks to recent advances in mathematics, a new self-learning approach is possible, and has been proven to address the problem of these new and advanced automated threats, at scale. This self-learning approach:

- Builds a sophisticated 'pattern of life' – understands what represents normality for every person and device
- Uses unsupervised technology – spots new threats as they emerge
- Does not depend on rules or signatures
- Enables continual internal monitoring

The 'immune system' approach

Applying the principles of a biological immune system to safeguard the health of the network, Darktrace's Enterprise Immune System is the world's first truly self-learning cyber defense solution.

Darktrace's machine learning technology, developed by mathematicians from the University of Cambridge, can detect previously unidentified threats, without rules, and automatically defend networks. Today's attacks can be of such severity and speed that a human response cannot happen quickly enough. Thanks to these self-learning advances, it is now possible for a machine to uncover emerging threats and deploy appropriate, real-time responses to fight back against the most serious cyber-threats.

"Darktrace's machine learning and mathematics are extremely powerful in detecting activity that is abnormal and will be critical to our future cyber security."

Mark Hughes, President of BT Security

Sophisticated Cyber-Threat and The Failure of Legacy Approaches

Today's cyber-threats are increasingly sophisticated, driven by intelligent humans or machines using unpredictable methods with time on their side. With an ever-growing number of connections, internally and externally, it has become increasingly difficult to track all network activity, and to set parameters and signatures that will provide sufficient levels of protection. The perimeters of networks have essentially become redundant, while cyber-threats are constantly evolving and developing new techniques.

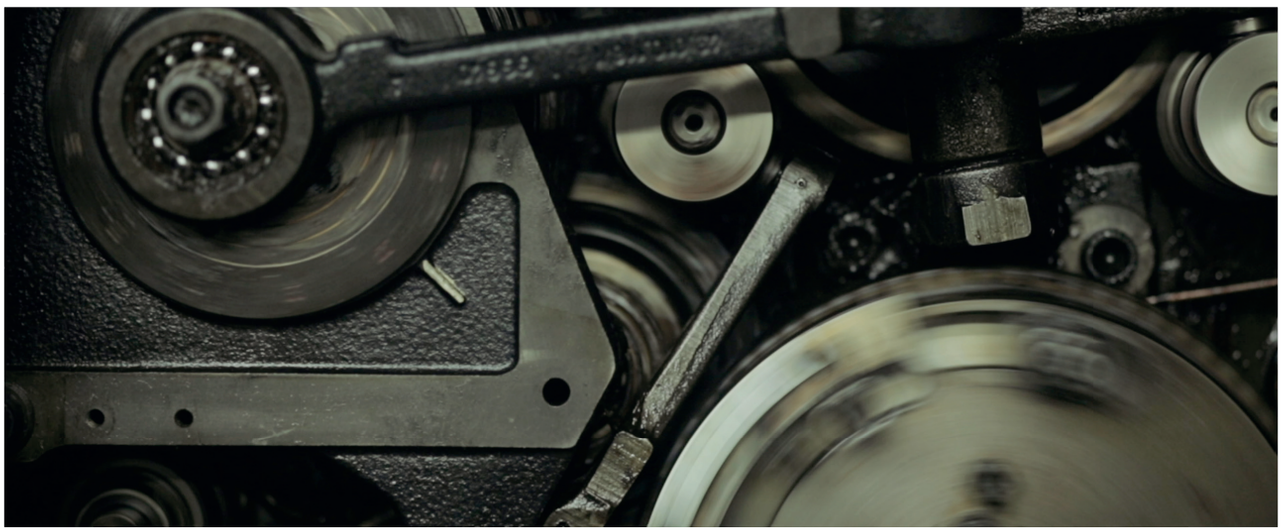
In the cyber security environment, firewalls, endpoint security methods and other tools such as SIEMs and sandboxes are deployed to enforce specific policies, and provide protection against certain threats. These tools form an important part of an organization's cyber defense strategy, but they are insufficient in the new age of cyber threat. Such tools have gradually become defunct or commoditized as networks grow, and advanced threats increasingly bypass these controls.

Ultimately, legacy systems have failed to contend with the developing cyber threat landscape. Approaches that depend on rules and signatures do not work:

- They need to know about *all* previous attacks.
- They need to perfectly understand your business and business-specific rules.
- They need a perfect way of sharing high quality information about new attacks.
- They need to *guess* what *all* future attacks and software weaknesses look like.
- They need to be able to turn *all* the above insight into rules or signatures that work.

For a cyber defense strategy dependent on rules and signatures to work, perfect knowledge of all past and future threats is required. This is not achievable.

- Border controls are dependent on signatures and recognition – if they miss an attack at the point of entry, they have failed and cannot take further action.
- Endpoint security depends on recognizing signatures and detecting attacks that have been previously identified – incapable of meeting the challenges of unknown threats.
- Intelligent sandboxes are defeated by more intelligent attacks, which recognize when they're in a fake space and delay the execution of malicious activity.
- Log tools and SIEM databases require an impossible level of manual effort to analyze each piece of data as it arrives, and they do not have learning capabilities.
- So-called 'behavioral analytics' cannot detect new and advanced threats as they emerge, as they are dependent on prior knowledge – signature-based approach in disguise.



Machine Learning and The History of Automation

Today, we are experiencing the third great revolution in automation: the era of machine learning. Recent advances in engineering and mathematics have made possible new classes of intelligent systems that are capable of making value judgements and carrying out higher value, more thoughtful tasks.

The history of automation dates back to the Industrial Revolution in the eighteenth and nineteenth centuries, where machines came to replace muscle power. It triggered a boom in industrial output and represented the birth of modern manufacturing. The introduction of machines saved time and energy, and freed humans to undertake other higher value tasks. They also greatly improved accuracy and precision.

The second revolution in the history of automation was the rise of computing. Computers started to automatically execute repetitive tasks that had previously been carried out by teams of people, such as payroll or stock taking. Thus, humans were liberated to focus on tasks that required considered thought and judgements. Using human capabilities in conjunction with computers increases productivity and quality.

In the age of information overload, these technological advancements are more significant than ever. As we wade through oceans of data, machine learning helps us clear the way by processing and making sense of that information. Machine learning is defining the industries of the next generation, especially true for the new field of cyber security.

"Darktrace's Enterprise Immune System has given us visibility into all our digital interactions. This ultimately means that our customers and data are better protected."

Philip Aim, Managing Director, CreaCard

Machine Learning To Date

The proliferation of data in the modern world means that it is not just unproductive, but impossible for humans to sift through the vast amount of information gathered each minute within a network.

Machine learning is difficult to develop and deliver, as it requires complex algorithms to be devised and an overarching framework to interpret the results produced. However, when applied correctly these approaches can facilitate machines to make logical, probability-based decisions and undertake thoughtful tasks.

We have already seen machine learning at work in a number of different commercial and industrial fields, for example:

- Payment processing companies can use state-of-the-art machine learning techniques to build models which can identify fraudulent payments in real time.
- Online video services use algorithms to understand the viewing preferences of customers in order to provide tailored programme recommendations for subscribers.
- Advertising firms are able to use analysis of browsing history to determine which adverts to make visible, making targeted decisions that deliver greater success than would otherwise be possible with human marketers.
- On-board computers in cars produce huge amounts of data which can be distilled to provide engineers with a better understanding of how customers actually use the vehicle, and also assist in the prediction of part failure.
- In healthcare, similar data collection processes mean that wellbeing can be closely monitored, problems highlighted earlier and therefore the risk of serious situations developing can be reduced.

Much of today's existing machine learning is supervised, however. This means that in order for the machine learning to operate successfully, there needs to be prior knowledge of the potential outcomes pre-programmed by a human. In industries where behaviors are well understood, these can prove more than adequate for assisting with product development or consumer safety.

However, in an area as complex and obfuscated as cyber security, there cannot be complete knowledge of all threats, existing and emerging.

The limitations of supervised machine learning in cyber security

Traditional approaches to cyber security are based on identifying activities that resemble previously known attacks – the 'known knowns'. This is usually done with a signature-based approach, whereby a database of known malicious behaviors is created. New activities are compared to those in the database and any which match are flagged as threats.

Other systems use methods based on supervised machine learning. Using this supervised approach, a system is trained using a data set in which each entry has been labelled as belonging to one of a set of distinct classes. In the information security context, the security system is trained using a database of previously seen behaviors, where each set of behaviors is known to be either malicious or benign and is labelled as such.

New activities are then analyzed to see whether they more closely resemble those in the malicious class or those in the benign class. Any that are evaluated as being sufficiently likely to be malicious are again flagged as threats.

Solely supervised machine learning approaches have fundamental weaknesses:

- Malicious behaviors that deviate sufficiently in character from those seen before will fail to be classified as such, hence will pass undetected.
- A large amount of human input is needed to label the training data.
- Any mislabelled data can seriously compromise the ability of the system to correctly classify new activities.

Unsupervised Machine Learning Applied to Cyber Defense

Machine learning presents a significant opportunity to the cyber security industry. New machine learning methods promise to enhance network visibility and improve detection levels thanks to the greater amount of computational analysis they can instigate.

Advanced machine learning is at the forefront of the fight against automated and human-driven cyber-threats, overcoming the limitations of rules and signature-based approaches:

- It learns what is normal within a network – it doesn't depend upon knowledge of previous attacks.
- It thrives on the scale, complexity and diversity of modern businesses, where every device and person is slightly different.
- It turns the innovation of attackers against them – any unusual activity is visible.
- It constantly revisits assumptions about behavior, using probabilistic mathematics.
- It is always up to date and not reliant on human input.

Utilizing machine learning in cyber security technology is difficult, but when correctly implemented it is extremely powerful. It means that previously unidentified threats can be detected, even when their manifestations fail to trigger any rule set or signature. Instead, machine learning allows the system to analyze large sets of data and learn a 'pattern of life' for what it sees.

Machine learning can attribute human capabilities to machines, such as:

- Thought: it uses past information and insights to form its judgements
- Real time: the system processes information as it goes
- Self-improving: its understanding is constantly being challenged and adapted, based on new information

New unsupervised machine learning therefore allows computers to recognize evolving threats, without prior warning or supervision.

Supervised vs Unsupervised Machine Learning

Supervised machine learning

Supervised learning is where a computer takes a series of data and sorts it according to given labels. For example, if we have a series of different animals, we can work out which of the animals are reptiles and which are mammals. We have defined the class, or label, and the computer sorts the data by discrimination.

Unsupervised machine learning

Unsupervised learning works things out without pre-defined labels. In the case of sorting the series of different animals, the system analyzes the information and works out the different classes of animals. This allows the system to handle the unexpected and embrace uncertainty. It does not always know what it is looking for, but can independently classify data and detect compelling patterns.

Darktrace's Unsupervised Machine Learning

Darktrace's unsupervised machine learning methods do not require training data with pre-defined labels. Instead they are able to identify key patterns and trends in the data, without the need for human input. The advantage of unsupervised learning is that it allows computers to go beyond what their programmers already know and discover previously unknown relationships.

Darktrace uses unique unsupervised machine learning algorithms to analyze network data at scale, intelligently handle the unexpected, and embrace uncertainty.

Instead of relying on knowledge of past threats to be able to know what to look for, it is able to independently classify data and detect compelling patterns that define what may be considered to be normal behavior. Any new behaviors that deviate from those, which constitute this notion of 'normality,' may indicate threat or compromise.

The impact of Darktrace's unsupervised machine learning on cyber security is transformative:

- Threats from within, which would otherwise go undetected, can be spotted, highlighted, contextually prioritized and isolated using these algorithms.
- The application of machine learning has the potential to provide total network visibility and far greater detection levels, ensuring that networks have an internal defense mechanism.
- It has the capability to learn when to action automatic responses against the most serious cyber threats.

"I intuitively feel that technology working as a self-learning immune system is the right way to do cyber defence."

Svein Ringbakken, Managing Director, DNK

Reverend Thomas Bayes

The cutting-edge mathematics at the forefront of Darktrace's machine learning approach are anchored in the seminal work of British mathematician Thomas Bayes (1702–1761). His theory of conditional probability provides a mathematical bridge between objective, developed methods and the subjective world that we populate. An advanced approach to Bayesian theory, developed by mathematicians from the University of Cambridge, provides a filter to ascertain the true meaning of messy and profuse data.

Darktrace's use of Bayesian probability as part of its unsupervised machine learning approach uniquely enables our technology to:

- Discover previously unknown relationships
- Independently classify data
- Detect compelling patterns that define what might be considered normal behavior
- Work without prior assumptions when needed

Borne out of Cambridge

As a world-class center of excellence for science and mathematics, the University of Cambridge is at the heart of new machine learning research. Cambridge is home to a hub of global technology leaders applying ground-breaking discoveries to a range of fields and industries.

Darktrace was founded under the supervision of two renowned individuals with a background in mathematics and machine learning: Professor Bill Fitzgerald and Dr Mike Lynch.

Professor Bill Fitzgerald was Professor of Applied Statistics and Signal Processing, and the Head of Research, in the Signal Processing Laboratory at the University of Cambridge. His ground-breaking work on Bayesian statistical methodology as applied to signal and data modelling had a profound impact on the study of signal processing, both within the University of Cambridge and internationally.

Dr Mike Lynch OBE is a renowned technologist, Fellow of the Royal Society, and an advisor to the UK government. His work and research in the area of Bayesian mathematics and machine learning built multi-billion dollar company Autonomy. With a PhD in Signal Processing from the University of Cambridge, Dr Lynch has a history of applying Bayesian mathematics approaches to real-world environments and infrastructures, and advises Darktrace, as the founder Invoke Capital.

Following fundamental developments at the University of Cambridge, statistics and signal processing are now diverse fields. These advances allow computers to extract small signals from large, high-dimensional data sets, often in real time.

This new mathematics not only identifies meaningful relationships within data, but also quantifies the uncertainty associated with such inference. By knowing and understanding this uncertainty, it becomes possible to bring together many results within a consistent framework – the basis of Bayesian probabilistic analysis.

The mathematics behind machine learning are extremely complex and difficult to get right. Robust, dependable algorithms need to be developed, with a scalability that enables their successful application to real-world environments. With leading scientists and mathematicians from Cambridge and other world-leading centers of education, Darktrace has created the first ever machine learning approach proven to work for cyber security at scale.

Technical Overview

A closer look at Darktrace's machine learning algorithms and approaches.

Darktrace's probabilistic approach to cyber security is based on a Bayesian framework. This allows it to integrate a huge number of weak indicators of potentially anomalous network behavior to produce a single clear measure of how likely a network device is to be compromised.

This probabilistic mathematical approach is critical to Darktrace's unique ability to understand important information, amid the noise of the network – even when it does not know what it is looking for.

Ranking threat

Crucially, Darktrace's approach accounts for the inevitable ambiguities that exist in data, and distinguishes between the subtly differing levels of evidence that different pieces of data may contain. Instead of generating the simple binary outputs 'malicious' or 'benign,' Darktrace's mathematical algorithms produce outputs that indicate differing degrees of potential compromise. This output enables users of the system to rank different alerts in a rigorous manner and prioritize those which most urgently require action, simultaneously removing the problem of numerous false positives associated with a rule-based approach.

At its core, Darktrace mathematically characterizes what constitutes 'normal' behavior based on the analysis of a large number of different measures of a devices network behavior:

- Server access
- Data volumes
- Timings of events
- Credential use
- DNS requests

Each measure of network behavior is then monitored in real time to detect anomalous behaviors.

Clustering

To be able to properly model what should be considered as normal for a device, its behavior must be analyzed in the context of other similar devices on the network. To accomplish this, Darktrace leverages the power of unsupervised learning to algorithmically identify naturally occurring groupings of devices, a task which is impossible to do manually on even modestly sized networks.

In order to achieve as holistic a view of the relationships within the network as possible, Darktrace simultaneously employs a number of different clustering methods including matrix based clustering, density based clustering and hierarchical clustering techniques. The resulting clusters are then used to inform the modelling of the normative behaviors of individual devices.

Clustering: At a glance

- Analyzes behavior in the context of other similar devices on the network
- Algorithms identify naturally occurring groupings of devices – impossible to do manually
- Simultaneously runs a number of different clustering methods to inform the models

Network topology

Any cyber threat detection system must also recognize that a network is far more than the sum of its individual parts, with much of its meaning contained in the relationships among its different entities, and that complex threats can often induce subtle changes in this network structure. To capture such threats, Darktrace employs several different mathematical methods in order to be able to model multiple facets of a networks topology.

One approach is based on iterative matrix methods that reveal important connectivity structures within the network in a similar way to that in which Google's PageRank algorithm reveals important relationships within the structure of the internet. In tandem with these, Darktrace has developed innovative applications of models from the field of statistical physics, which allow the modelling of a network's 'energy landscape' to reveal anomalous substructures that may be concealed within.

Network structure

A further important challenge in modelling the behaviors of network devices, as well as of networks themselves, is the high-dimensional structure of the problem with the existence of a huge number of potential predictor variables. Observing packet traffic and host activity within an enterprise LAN or WAN is difficult because both input and output can contain many inter-related features (protocols, source and destination machines, log changes and rule triggers, etc.). Learning a sparse and consistent structured predictive function is crucial to avoid the curse of overfitting.

In this context, Darktrace has employed a cutting edge large-scale computational approach to learn sparse structure in models of network behavior and connectivity based on applying L1-regularization techniques (the lasso method). This allows for the discovery of true associations between different network components and events which can be cast as efficiently solvable convex optimization problems and yield parsimonious models.

Recursive Bayesian Estimation

To combine these multiple analyses of different measures of network behavior to generate a single comprehensive picture of the state of each device, Darktrace takes advantage of the power of Recursive Bayesian Estimation (RBE) via a novel implementation of the Bayes filter. Using RBE, Darktrace's mathematical models are able to constantly adapt themselves, in a computationally efficient manner, as new information becomes available to the system. They continually recalculate threat levels in the light of new data, identifying changing attack behaviors where conventional signature-based methods fall down.

Darktrace's innovative approach to cyber security has pioneered the use of Bayesian methods for tracking changing device behaviors and computer network structures. The core of Darktrace's mathematical modelling is the determination of normative behavior, enabled by a sophisticated software platform that allows for its mathematical models to be applied to new network data in real time.

The result is a system that is able to identify subtle variations in machine events within a computer networks behavioral history that may indicate cyber-threat or compromise.

$$\tilde{p}(x_t|y_t) = \sum_{i=1}^{N} w^{(i)} \times \delta\left(x_t^{(i)}\right)$$

"Darktrace is the only company that uses mathematical analysis and machine learning to detect potential threats, allowing us to stay ahead of evolving risks."

Mike Somers, Information Systems and Security Manager, Open Energi

Conclusion – A New Approach to Cyber Security

Our generation is witnessing the machine learning revolution – the third great era of automation. We are seeing shifts in working practices brought about by the replacement of muscle with machine, the automation of repetitive tasks, and now the replacement of low value, thoughtful tasks with machines capable of handling big data and making vast calculations.

As networks have grown in scope and complexity, the opportunities for attackers to exploit the gaps have increased. Walls are no longer enough to protect the content of systems; rules cannot pre-emptively defend against all possible attack vectors; signature-based detection methods fail repeatedly. Cyber attacks are advanced, subtle and varied – only automated responses based on machine learning can keep pace with them.

Machine learning is difficult to deliver, but Darktrace is proving that it works. By utilizing the probabilistic Bayesian mathematics developed by mathematicians from the University of Cambridge, Darktrace is at the forefront of machine learning advances. Our unsupervised detection methods are the spearhead of the defensive response, demonstrating the power that unsupervised machine learning can have.

The Enterprise Immune System approach means that detection no longer depends on an archive of previous attacks. Instead, attacks can be spotted against the background understanding of what represents normality within a network. No pre-definitions are needed, which allows for the best possible insight and defense against today's threats.

On top of the detection capability, the Enterprise Immune System can create digital antibodies automatically, as an immediate response to the most threatening cyber breaches. The immune system approach both detects and defends against cyber threat.

Genuine unsupervised machine learning eliminates the dependence on signature-based approaches to cyber security, which are not working. Darktrace's technology has become a vital tool for security teams attempting to understand the scale of their network, observe levels of activity, and detect areas of potential weakness. These no longer need to be manually sought out, but are flagged by the automated system and ranked in terms of their significance.

Machine learning technology is the fundamental ally in the defense of systems from the hackers and insider threats of today, and in formulating response to unknown methods of cyber attack. It is a momentous step change in cyber security. Defense must start within.

About Darktrace

Named 'Technology Pioneer' by the World Economic Forum, Darktrace is one of the world's leading cyber threat defense companies. Its Enterprise Immune System technology detects previously unidentified threats in real time, powered by machine learning and mathematics developed at the University of Cambridge, which analyze the behavior of every device, user and network within an organization. Some of the world's largest corporations rely on Darktrace's self-learning appliance in sectors including energy and utilities, financial services, telecommunications, healthcare, manufacturing, retail and transportation. The company was founded in 2013 by leading machine learning specialists and government intelligence experts, and is headquartered in Cambridge, UK and San Francisco, including offices in Auckland, Boston, Chicago, Dallas, London, Los Angeles, Milan, Mumbai, New York, Paris, Seoul, Singapore, Sydney, Tokyo, Toronto and Washington D.C.

Contact Us

US: +1 (917) 363 0822

Europe: +44 (0) 1223 350 653

Email: info@darktrace.com

www.darktrace.com