WHITE PAPER

Big Data and Predictive Analytics: On the Cybersecurity Front Line

Sponsored by: SAS

Robert Eastman
Alan Webber
Michael Versace
February 2015

IDC OPINION

The Challenge

"The simple fact is that we have a target on our chest. The attacks are going to come. We have to be agile in our response, and a key part of that is having the right information at the right time to be able to respond in the right way." (Government CIO)

The cybersecurity environment is shifting. Firms and government agencies are faced with larger and broader changes to their industries and organizations than ever before. Most of this change comes from the ways we create, collect, store, and share information. This information, which is such a tempting target for hackers, criminals, nation-states, and others, makes cybersecurity one of the top issues on every executive's mind.

IDC interviewed information security executives, practitioners, and industry experts across three industries: the U.S. federal government, financial services, and energy. The purpose of the study was to better understand the evolving cybersecurity threat landscape and the role that the emerging technologies of Big Data and predictive analytics would play in mitigating the threats and risks these industries face every day. Through these interviews and additional background research, IDC confirmed that the cybersecurity threats facing these industries are forcing business and agency security managers to rethink and reevaluate the tools and tactics they deploy to battle the cybersecurity threat. IDC's research uncovered that cybersecurity threats are evolving rapidly and that firms and government agencies must shift from a reactive approach to a proactive approach — understanding the threat before an attacker can cause damage. Shifting to a proactive approach requires organizations to tap into all available information and apply predictive and behavioral analytic tools to discover the potential of a threat, detect the actual threat, gather intelligence about the attack, and execute an enterprisewide response before the threat becomes significant.

A proactive posture includes advanced threat detection, real-time identification of risks, and protection and countermeasures to ensure that the increasing number and variations of cyberattacks are identified and mitigated before they have material financial or reputational impact. Organizational executives are recognizing that getting there requires a different tool set — a combination of Big Data and predictive analytics, used in ways not envisioned before.

IN THIS WHITE PAPER

In this white paper, IDC discusses the shifts happening in cybersecurity through the introduction of Big Data and analytics into the cybersecurity tool set. Chief information security officers (CISOs) and chief risk officers (CROs) have a number of different tools to improve cybersecurity and reduce risk, but the introduction of Big Data as a source for organizational intelligence creates both a challenge and an opportunity. Applying behavioral analytics to enterprise and external data assets to evaluate both internal and external threats gives CISOs and CROs a whole new level of insight and an improved ability to respond.

SITUATION OVERVIEW

The Rise of Big Data and Analytics

Every organization is being overwhelmed by data. Organizational data isn't new — organizations have been collecting data about accounting and finance, operations, supply chain, and customers for a long time. What is new, and what makes Big Data different, is the technology that gives a business the ability to collect and use this information in new ways.

First, the rate of data generation has increased through both the ability to capture previously uncapturable data and new sources of data. The growth in sources of data is extremely broad — including social media, facility operations data, geospatial data, threat intelligence, SQL server logs, security information and event management (SIEM) data, NetFlow data, firewall logs, intrusion detection system (IDS) data, Active Directory files, and more. Much of this data has always been there, but organizations weren't able to capture it until newer methods and technologies were developed and put in place.

Second, growth in data sources has resulted in an increase in the volume of data that organizations have to deal with. Thanks to dropping storage costs, either in-house or through a cloud-based solution, all of this data that was once retained only for a finite time is now retained in growing data sets and data filing systems that can be held indefinitely.

Third, the rate of data generation and the increasing number of sources have resulted in a larger variation in the types of data, ranging from the highly structured data set to the highly unstructured data set. Historically, most data available for analysis was primarily numerical in format. Today, the types of data being captured range from transaction-based internal financial accounting data to structured customer relationship data to highly unstructured data drawn from email, social media, threat feeds, and data sources that are being created through a combination of existing data points.

Big Data was once the purview of a few online firms and government agencies; now, the avalanche of Big Data is burying most organizations. The issue is not the types of data, the gathering of data, or the storage of data – the issue comes down to what organizations do with the data. To drive value from Big Data, organizations turn to behavioral analytics and frameworks such as Hadoop to make better decisions, minimize costs, maximize return on investments, and optimize operations. Information security can no longer be overlooked, and CISOs and CROs are most interested in using Big Data and, correspondingly, analytics that will improve their organization's security posture.

The Cybersecurity Threat Landscape

The world of cybersecurity is changing at a rapid pace. IDC has found that the number of attacks continues to rise, the types of threat actors have increased, and the attack vectors have grown and migrated to more sophisticated advanced persistent threats (APTs), insider attacks, fraud, and cybercrimes. Traditional security solutions and mindsets are not sufficient to deal with this new threat landscape.

The concept of a cybersecurity attack or incident requires a broad definition. These incidents can range from the misplacement of hardcopy material that contains personally identifiable information (PII) that can lead to a data breach to much more long-term and damaging APTs. Most current cybersecurity threats can be categorized into the following broad categories:

- Advanced persistent threats
- Distributed denial of service (DDoS)
- External software introduction including malware and adware
- Insider data theft
- Phishing, spear phishing, and other forms of email-based spoofing and fraud
- Social engineering and other forms of physiological manipulation
- SQL injection and other code injection techniques
- Trojan attacks
- URL redirection or parameter tampering
- Zero-day attacks

Though actors may come from a broad set of backgrounds with varying degrees of financial support and motivators, most threat actors can be categorized into one of six categories:

- Accidental: Generally an insider such as an employee or a contractor; causes harm accidently because of inexperience
- Insider: Often inexperienced but can have higher-level skills; uses opportunities to target known vulnerabilities in systems and policies for self-gain

"What we have now isn't good enough for tomorrow and maybe not even today. Government CIOs and CISOs need new technologies to address the ever-evolving threat we face." (Government IT Security Manager)

- **Opportunist:** An external party who lacks significant experience but uses opportunities to target known vulnerabilities employing worms, viruses, bits, and other tools; often done for bragging rights
- **Hacktivist:** External party with higher-level skills that target known vulnerabilities using DDoS attacks or malware as a path to introducing more sophisticated tools into a target system; often has a political or similar motive for action
- **Professional criminal:** Organized crime efforts including terrorist groups that use high-level and sophisticated skills to target financially relevant information
- **State-level actor:** Generally state-level actors or those who work on behalf of a national government — including industrial espionage — using high-level and sophisticated skills to target strategic or economic information

The targets and vectors for attacks also continue to evolve with new technologies. Though databases and Web sites remain traditional targets, IDC sees threats and vulnerabilities in several new areas, including attacks on:

- Social media and mobile devices
- Employee-owned devices (bring your own device [BYOD])
- Private, hybrid, or public clouds
- The Internet of Things (IoT) where a wide variety of devices are connected to the Internet

The propagation of attack vectors and threat actors against an increasing set of target areas has resulted in an exponentially growing level of cybersecurity complexity for the CISO and the CIO to deal with. The result is that proactive mitigation of these threats with current technologies is almost impossible unless a new approach is applied.

Intersection of Big Data, Analytics, and Cybersecurity

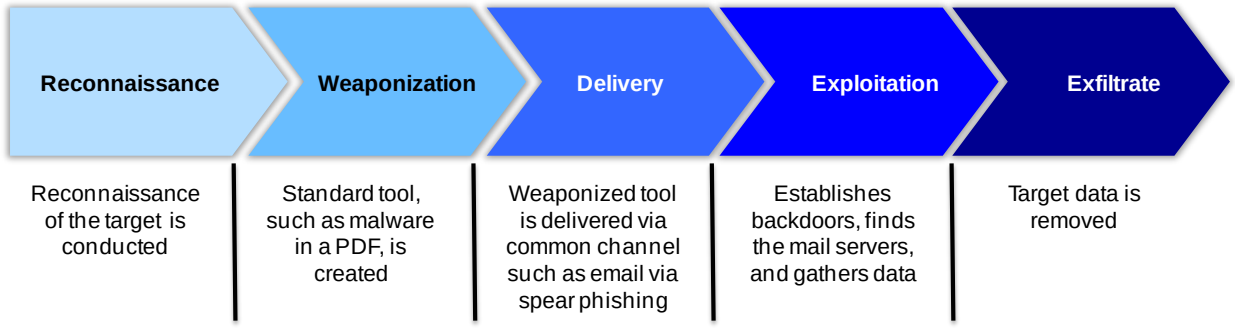
Compared with traditional cybersecurity methods and efforts, Big Data and behavioral analytics offer the opportunity to improve situational awareness and information security, or what IDC calls cyberanalytics. Through appropriately timed cyberanalytics, context can be provided so that patterns and anomalous behaviors can be identified that are indicators of fraud, theft, or other security breach.

Information security that is data driven is not a new concept. For example, in the financial services vertical, Visa, MasterCard, and American Express have used data and analytics to detect potentially fraudulent transactions based upon patterns and pattern recognition across millions of transactions. In government, agencies have been using data and analytics to discover terrorist threats and fraud in social programs and detect insider threats, as well as in other intelligence applications. Today, the limitation for cybersecurity is the ability for solutions and organizations to leverage all data assets. The historical dump-and-analyze approach has proven ineffective because the needed data history is not typically stored or analyzed in a timely fashion. New approaches are required to leverage and evaluate data in a way that can determine what data is and isn't important to support cybersecurity-based behavioral analytics.

Currently, CISOs in most organizations use analytics to model known threats and vectors. But because of changing attack vectors and methods, the actual time to detect and respond to a threat has increased. Next-generation cyberanalytics now has the capability of situational awareness on both internal and external threats to an organization through modeling for both good (normal) and bad (anomalous) behaviors. For example, previously it was difficult to detect and expose an insider threat until long after the damage had been done because of the normal path of an attack (see Figure 1). Now, through the use of cyberanalytics, contextualization, and correlation, companies can predict unusual behavior and detect an active insider threat by anomaly detection in interaction behavior such as accessing a database and downloading a set of files. Unless an organization has the ability to model normal and anomalous behavior of people and network assets, it will be unable to detect these newer types of attacks.

FIGURE 1

Anatomy of an Advanced Persistent Threat Attack



Source: IDC, 2015

Big Data, Analytics, and Cybersecurity — An Industry Perspective

Government

IT security is not a small or inexpensive problem for government. According to the United States Computer Emergency Readiness Team (US-CERT), over 46,000 cyberincidents occurred at all U.S. federal government agencies in 2013. Incidents included loss of PII, Web sites being hacked and defaced, and penetration of secure databases. IDC estimates that to thwart attackers and address the number of incidents, U.S. federal government agencies alone will spend over $14.5 billion on IT security.

Government CISOs face a unique challenge – the infrastructure and systems that they are charged with protecting often range from older mainframe systems to core ERP systems to newer cloud-based databases and even mobile apps. Any set of security solutions needs to stretch across the breadth of the different systems and applications that government runs.

Government agencies and databases are prime targets for all types of attacks. Past attacks on government systems, databases, and Web sites have included APTs, DDoS, SQL injection attacks, Trojans, and malware. The sources of those attacks have included clumsy insiders as well as foreign nation-states. This increasingly hostile environment is a top concern for agency and department heads and has pushed government toward the adoption and implementation of more complex cybersecurity tools, including predictive behavioral analytics.

"The amount of data that we now have is overwhelming from so many sources we can't keep track. The key is how do we get security value from the data." (Government IT Manager)

The most prominent threats that government agencies face include:

- **BYOD.** Smartphones and tablets are wonderful technologies. With cameras, microphones, large processing power, and the ability to function outside of normal networks, they are also a threat. Government employees and contractors are increasingly using their own mobile devices during work and even on government systems. By doing so, they are potentially opening government systems to attacks.
- **Internal agents.** Internal threats are nothing new to government but have been given more exposure in light of recent events over the past few years. The introduction of more and more technology into the government workplace will only increase the number of internal threats that government faces.
- **Social engineering.** Whether phishing, spear phishing, or other social engineering technique, government is susceptible to social engineering attacks that can result in the placement of malware or Trojans on systems and lead to an APT attack.
- **Targeted malware.** Malware is getting smarter. New forms of malware are being developed that — based upon previous defense efforts — can be targeted toward only specific systems and that cannot be executed on any other systems. This is especially critical as government systems become more of a target.
- **APTs.** Highly sophisticated and carefully planned, APTs are a significant risk to government systems. Meant to penetrate government systems and then quietly steal information over time, APTs focus on higher-value information.

As government agencies shift from a reactive defensive posture to a proactive posture through the adoption of Big Data and predictive analytics technology, there are several key challenges that must be considered (see Table 1). In overcoming these challenges, agencies can achieve comprehensive visibility into potential risks and active threats and define mitigation plans in a reduced time frame.

TABLE 1

Significant Cybersecurity Challenges in Government

Challenge	Implications
Complicated system architecture	Agencies are forced to deploy overly complicated, multinode, multilayered security solutions to match the deployments of technologies that range from mainframe computers to smartphones, with each node being a potential attack vector.
Internet of Things (IoT) technologies	The Internet of Things as deployed by government agencies, such as with smart meters, video cameras, and other technologies, will provide additional sources of data and additional vulnerability nodes.
Privacy	The ability of government agencies to collect information about citizens will continue to grow as the amount of information about citizens grows. These databases of citizen information will become prime targets for attackers at the same time government tries to establish rules for collection and protection of private information.

Source: IDC, 2015

Financial Services

Evolving scope of cyberattacks on financial enterprises, coupled with industry reputation and increased regulatory pressures, is the prime motivator for cybersecurity investments. Cybersecurity strategies remain at the top of the agenda as policymakers and board-level executives around the globe continue to focus on capital buffers, trade transparency, reputational risks, financial crime and fraud, and the impact of customer losses on the safety and soundness of the financial marketplace. Worldwide, of the $75.4 billion spent by the financial services industry on technologies and services for risk management, over $40.0 billion will be spent on managing operational risks including cyberrisk, with $27.4 billion spent specifically on information security and fraud.

At the center of cybersecurity requirements are the increased sophistication of fraud actors, shrinking response windows, and the complexity of the threats to digital channels. The result is a growing number of requirements that have made advanced, predictive threat intelligence solutions and services top agenda items for chief risk officers, data officers, executives, and regulators.

"CISOs must enable the businesses they support to make more informed decisions … many times when their businesses only care about getting things done. The challenge is shifting from being the 'no' police to an enabler that makes it safer for the firm to operate." (CISO, Global Hedge Fund)

The most prominent threats that financial institutions face include:

- **Third-party vulnerability.** Financial institutions, as part of a large and complex financial network, partner with a number of other firms, from credit card firms to payment processors to clearinghouses and more. Many of these third-party providers may not have the same level of security and protection as a financial institution and have the potential for introducing unknown risks into the relationship.
- **Wireless payment systems.** More and more vendors, payment processors, and consumers are adopting wireless payment systems such as Apple Pay. These new wireless payment systems open up a whole new set of potential targets for attackers.
- **Cyberattacks go global.** Phishing, targeted malware, ATM skimming, and other threats targeted at financial institutions used to be primarily a first-world problem. But the global growth in online banking and digital interactions has introduced these problems into other parts of the world, and it is expected that as adoption rates increase, so will the spread of these problems.
- **Targeted social engineering.** Whether phishing, spear phishing, pharming, botnets/zombies, or other technique, financial institutions face an avalanche of attacks that target account takeovers, network disruption, and identity theft.

As part of a future security infrastructure strategy, firms today must consider behavioral analytics-based intelligence solutions that are designed to provide visibility, insight, and recommendations from Big Data for intelligent security action within ever-shrinking time windows. At the limit, firms must design these analytic solutions with the ability to identify potential actors, breaches, and fraud before a transaction is processed. Behavioral analytics for security operations must also include the assessment of ongoing risk in meeting commitments and complying with policies — in both the short term and the long term — as part of planning processes and strategy formation.

IDC predicts that by 2020, Big Data, analytics, and cloud service operations will be the preferred delivery model for 50% of all cybersecurity requirements as financial institutions move aggressively to retire in-house legacy information security strategies and systems and deploy cyberanalytics capabilities. In considering deployment options, CIOs and CISOs must define a strategy that addresses the top 4 challenges: information governance, hybrid IT, shift to digital, and lack of a viable security workforce (see Table 2).

TABLE 2

Significant Cybersecurity Challenges in Financial Services

Challenge	Implication
Information governance and the growth in data and information risk	Growth in data is the fuel for security threats, fraud, and loss. More digital information, from customer data to intellectual property, is stored in insecure areas and beyond necessary retention time frames. Data is often exchanged between consumers and businesses, with few controls across the supply chain. Data is lost accidentally or as a result of poorly established control monitoring systems. IDC Financial Insights estimates that over the past two years, account information from over 450 million customers has been stolen or lost due to porous systems. Information governance programs are strained in their ability to scale up to rising data volumes and scale out to growth data, analytic modeling, and software types.
Hybrid IT	Virtualization of all datacenter assets, including compute, storage, database, network, and systems management, and the continued move toward ITO and BPO require a new baseline of security standards, risk sharing, and operational risk management capabilities. Hardware, software, services, and internal resources must be reoriented and adapted to an elastic computing environment where IT assets are completely distributed and movable across on-premises and external environments. Legacy security strategies break when moving to hybrid IT environments.
Digital financial services	Enterprise and consumer mobility is the backbone of digital financial services and creates new challenges for CISOs to address the growing number of devices and apps that connect to corporate digital channels.
Lack of security talent	The 3rd Platform of IT (cloud, Big Data/analytics, mobile, social) will change the demands on 95% of all IT skills, including security, over the next three to five years. This ultimately changes the balance between in-house skills and skills acquired through service providers.

Source: IDC, 2015

Energy and Utility

The utility industry, one of the most asset-intensive industries, is a rich target for cyberthreats and cyberattacks. Therefore, cybersecurity and risk management remain high on the utility CIO's agenda, specifically around protecting the utility's cyberassets and other assets, prevention of attacks, and compliance with cybersecurity-related regulatory mandates. Many utility sector systems such as supervisory control and data acquisition (SCADA) have been in use for some time, and few of these systems were designed with foreknowledge of the current cyberthreat environment. As network technologies have become more standardized, gaining access to networks has become easier, cyberattack strategies have become more sophisticated, and the threat level has increased.

In response, the North American Electric Reliability Corporation (NERC), a regulatory authority, is adopting and enforcing increasingly stringent standards for protection of critical assets in the industry.

The expanding series of NERC CIP (Critical Infrastructure Protection) standards, in particular, are forcing utilities to make significant investments in cybersecurity resources. The NERC CIP V5 standard has raised the stakes for what utilities must do to protect their critical infrastructure. But these efforts may not be enough.

For example, the Stuxnet virus in 2010 raised awareness of the vulnerability of process control systems. The Federal Energy Regulatory Commission (FERC) stated that an attack on less than a dozen of the tens of thousands of electric substations in the United States, under the right conditions, could cause a massive failure of the electrical grid, evoking memories of the 2003 blackout in the Northeast that left 50 million people without power. Incidents like this and others demonstrated that a limited attack on the right resources at the same time could cause widespread infrastructure damage.

The most prominent threats that the energy and utility industry faces include:

- **BYOD.** Smartphones and tablets are almost ubiquitous technologies and also a threat. Energy and utility industry employees and contractors are increasingly using their own mobile devices during work and potentially opening up systems to risks.
- **Social engineering.** Utilities are susceptible to social engineering attacks where people are seen as the most vulnerable link, which can result in the placement of malware or Trojans on systems, leading to an APT attack.
- **Targeted malware.** Malware is getting smarter. New forms of utilities-specific malware, such as Stuxnet, are being developed that – based upon previous defense efforts – can be targeted toward only specific utilities systems and that cannot be executed on any other systems. This is especially critical because utility systems are older and more vulnerable to known exploits. Though malware, such as Stuxnet, can be tailored, it was developed to compromise programmable logic controllers (PLCs) that control machinery and was originally targeted at Iran's nuclear centrifuges.
- **APTs.** Highly sophisticated and carefully planned, APTs are a significant risk to energy and utility systems. For example, the Dragonfly/Energetic Bear/Havex/Crouching Yeti incident in 2013 is an example of a sophisticated and multifaceted campaign making up an APT targeting the energy industry. In early 2013, the perpetrators started with phishing emails to target companies to tap the low-hanging fruit of human points of entry. The campaign then utilized watering hole attacks to redirect Web site visitors to a compromised alternate Web site designed to gain entry to the visitors' systems. Finally, the campaign infected the software bundles of three ICS equipment manufacturers with a remote access Trojan.

"All utilities receive noise. The threats that we worry about in the electric sector are different. What has changed over the past couple of years is that the impact of the threats is growing. This requires a more mature approach to the problem." (Utilities CISO)

Given the complexity and changing nature of threats, and the strict (and evolving) regulatory cybersecurity mandates, advanced and predictive analytics solutions offer important capabilities. Utilities are just beginning to appreciate the opportunities for greater threat identification and remediation that Big Data analytics can deliver in the area of cybersecurity. Utilities are realizing that they need vastly different tools today than just a few years ago and that their needs for advanced analytics are evolving as the nature of threats changes and regulatory mandates continue to develop.

As the utility industry grapples with the changing face of cybersecurity, it must address several key challenges while balancing competing investment requirements and severe resource constraints (see Table 3).

TABLE 3

Significant Cybersecurity Challenges in Energy and Utility

Challenge	Implication
Complex utility landscape	Utilities are complex, asset-rich environments that provide a broad surface for cybersecurity attacks and incidents. The number and breadth of assets that a utility has stretch the capability of the utility staff to secure each at-risk asset, even if many of these assets are not remotely located.
Criticality of assets	The utility industry has been deemed a critical infrastructure sector. By its very nature, the utility industry has attracted more than its share of cybersecurity threats.
Legacy assets	The utility industry has a considerable number of legacy systems and assets that were designed and deployed prior to the advent of the Internet. These assets are now exposed to a network, wittingly and unwittingly, and the attendant security capabilities and protections are not always current and state of the art, exposing these assets to security risks.

Source: IDC, 2015

CONCLUSION

"Business is all about taking risks, but you don't want to take on risks you don't know about." (CRO, U.S. Cyberinsurer)

Big Data and Analytics Join the Cybersecurity Front Line

Security, and specifically cybersecurity, is often rated as one of the top concerns for managers and leaders in both private sector and government organizations. Right now, new and evolving technologies – smartphones and IoT nodes as well as core systems and applications – are changing the ways organizations operate and interact with employees, customers, and citizens. These new technologies will enable new services and more effective and efficient operations, but they will also change the security landscape as they open up new vulnerabilities and provide new avenues for attackers. Organizations need a new set of security solutions that are flexible enough to adapt to changing technologies and are able to match the sophistication of the attacks they face.

Behavioral analytics has become a bet-the-business activity for many CROs; however, providing technological support for this specialized form of analytics presents significant challenges for security operations, including:

- **Scalability.** Behavioral analytics requires continuous, rapid ingestion of raw, granular data from multiple sources, resulting in ever-growing data volumes.
- **Optimization.** Rapid processing of analytical models at the point of a decision yields an optimized response. Traditionally, this has meant a highly custom approach to integrate multiple technology components in support of this demanding workload.
- **Expertise.** Beyond the technology itself, the biggest issue faced by institutions is the lack of skills to configure and tune analytic engines, interpret the insights generated, and act according to the potential threat.
- **Integration.** Cloud computing provides significant advantages for highly specialized, data-intensive analytic workloads that commonly define operational risk systems, requiring some degree of cloud and on-premises security operations. Firms must carefully define their security and operational strategies to strike an effective balance between managed and on-premise coverage, integration, and governance needs.

All organizations will continue to struggle with how to integrate solutions against an upward spiral of cyberattacks that target information for financial, reputational, and political gains. To really improve organizational security and reduce risk, organizations should:

- **Map and tap into existing data sources.** Organizations have a wealth of existing or easy-to-access data that could support improved security. The data sources range from network information to Web activity logs to NetFlow information to server logs and more. But organizations need to map these data sources and figure out what is already available, what is easily accessible, and what data is buried deeper but worth the effort. Once the data has been mapped, it then needs to be evaluated for its value in further analysis.
- **Contextualize and connect the data.** Once an organization understands the data it has, the organization needs to establish the context of the data and then correlate the data with other data, network information, and architectural components. This builds the foundational picture of the active risks and threats to the organization. Analytics can then be layered onto the correlated data to provide nuance to the picture.
- **Use deep analytics to refine and sharpen the picture.** Understanding the context of the data and then connecting the pieces of data is just the beginning. Using some of the newer analytics capabilities sharpens the picture through the operationalization of models, providing useful metrics, adequate information, and recommendations for decision making for current, ongoing, and future risks.
- **Move from reactive to proactive to real time.** Once analytics are in place and in use, analytics needs to move to real-time evaluation such that threats can be proactively mitigated before significant loss occurs. Analytics can also be done more holistically to detect "slow and low" threats that emerge over time.

As the cybersecurity threat increases and evolves, progressive organizations are realizing that one of their strongest resources to fight this threat lies in the growing volume of data at their disposal — and the increasing power of the technologies to act on this data. As enterprises accumulate more data — and more information — about what is going on inside their environments, this data holds vital behavioral clues for identifying both internal and external threats and risks. With the growing capabilities of advanced and predictive analytics, enterprises now have better technological means to identify and respond to the changing threat landscape. While the threats are unlikely to abate anytime soon, the new situational awareness that advanced and predictive analytical tools can deliver means that enterprises have an important technology ally on their side.

About IDC

International Data Corporation (IDC) is the premier global provider of market intelligence, advisory services, and events for the information technology, telecommunications and consumer technology markets. IDC helps IT professionals, business executives, and the investment community make fact-based decisions on technology purchases and business strategy. More than 1,100 IDC analysts provide global, regional, and local expertise on technology and industry opportunities and trends in over 110 countries worldwide. For 50 years, IDC has provided strategic insights to help our clients achieve their key business objectives. IDC is a subsidiary of IDG, the world's leading technology media, research, and events company.

Global Headquarters

5 Speen Street
Framingham, MA 01701
USA
508.872.8200
Twitter: @IDC
idc-insights-community.com
www.idc.com
