FROM TRADITIONAL ANTIVIRUS TO COLLECTIVE INTELLIGENCE

pandasecurity.com

01 Abstract
There is more malware than ever being released in the wild, and antivirus companies relying on signatures to protect users cannot keep up with the pace of creating signatures fast enough.

02 The malware landscape
Antivirus laboratories under attack. Malware techniques and design.

03 Panda’s Technology Evolution
First Generation: Antivirus. Second Generation: Anti-malware. Third Generation: Proactive technologies. Fourth Generation: Collective Intelligence.

04 Conclusion
The latest advances by the black hat and cybercrime communities are taking advantage of the inherent weaknesses in the security industry.

05 References

01 Abstract

There is more malware than ever being released in the wild, and antivirus companies relying on signatures to protect users cannot keep up with the pace of creating signatures fast enough.

As a result, the current installed base of anti-malware solutions is proving to be much less efective against the vast amounts of threats in circulation.

As we have been able to prove in a recent research study, even users protected with anti-malware and security solutions with the latest signature database are infected by active malware. Complementary approaches and technologies must be developed and implemented in order to raise the effectiveness to adequate levels.

This paper presents the fourth generation of security technology by Panda Security, called Collective Intelligence. Panda's Collective Intelligence allows us to maximize our malware detection capacity while at the same time minimizing the resource and bandwidth consumption of protected systems.

Collective Intelligence represents an approach to security radically different to current models. This approach is based on an exhaustive remate, centralized and real-time knowledge about malware and non-malicious applications maintained through the automatic processing of ali elements scanned.

One of the benefits of this approach is the automation of the entire malware detection and protection cycle (collection, analysis, classification and remediation). However, automation in and by itself is not enough to tackle the malware cat-and-mouse game. With large volumes of malware also comes targeted attacks and response time in these scenarios cannot be handled by automation of signature files alone.

The other main benefit that Collective Intelligence provides is that it allows us to gain visibility and knowledge into the processes running on all the computers that it scans. This visibility of the community, in addition to automation, is what allows us to tackle not only the large volumes of new malware but also targeted attacks.

02 The malware landscape

It is a known fact by all security professionals that there are more malware samples infecting users than ever before.

Malware writers have realized they can gain large amounts of money from distributing malware. The shift in motivation for creating malware, combined with the use of advanced techniques, has resulted in an exponential growth of criminally professional malware being created and distributed to infect unsuspecting users.

Also known as a type of targeted attacks, this new malware dynamic has become the next big plague for users and companies alike. Gartner estimates that by the end of this year 75% of enterprises will be infected with undetected, financially motivated, targeted malware that evaded their traditional perimeter and host defenses[2].

Antivirus laboratories under attack

Nowadays antivirus laboratories are under a constant and increasingly frequent distributed denial of service attack. The security industry is literally being saturated with thousands of new malware samples every day.

Each one of these new samples needs to be looked at by an analyst trained in reverse engineering in order to create a signature, which is costly and resource intensive from a corporate and business perspective.

Some companies are trying to deal with the problem by increasing the number of analysts in their labs[3] or by advocating for stronger intervention[4] by law enforcement[5] agencies by convicting the most active malware creators and easing the workload.

Initiatives to get law enforcement more involved are definitely a necessary step in the right direction. But unfortunately it seems an insufficient solution for the short term as the number of variants continues to increase and most cases, only the "mules" and "script kiddies" are actually convicted.

The more advanced malware writers, who are selling their code to spammers, mafias and criminals, are more evasive and harder to catch. In addition, the lack of resources at most law enforcement agencies around the world, tied to insufficient international cooperation and coordination among them make for a difficult task when trying to arrest a suspect or known cyber criminal. In the long run, both a technological and a social approach are needed if we want to solve this problem.

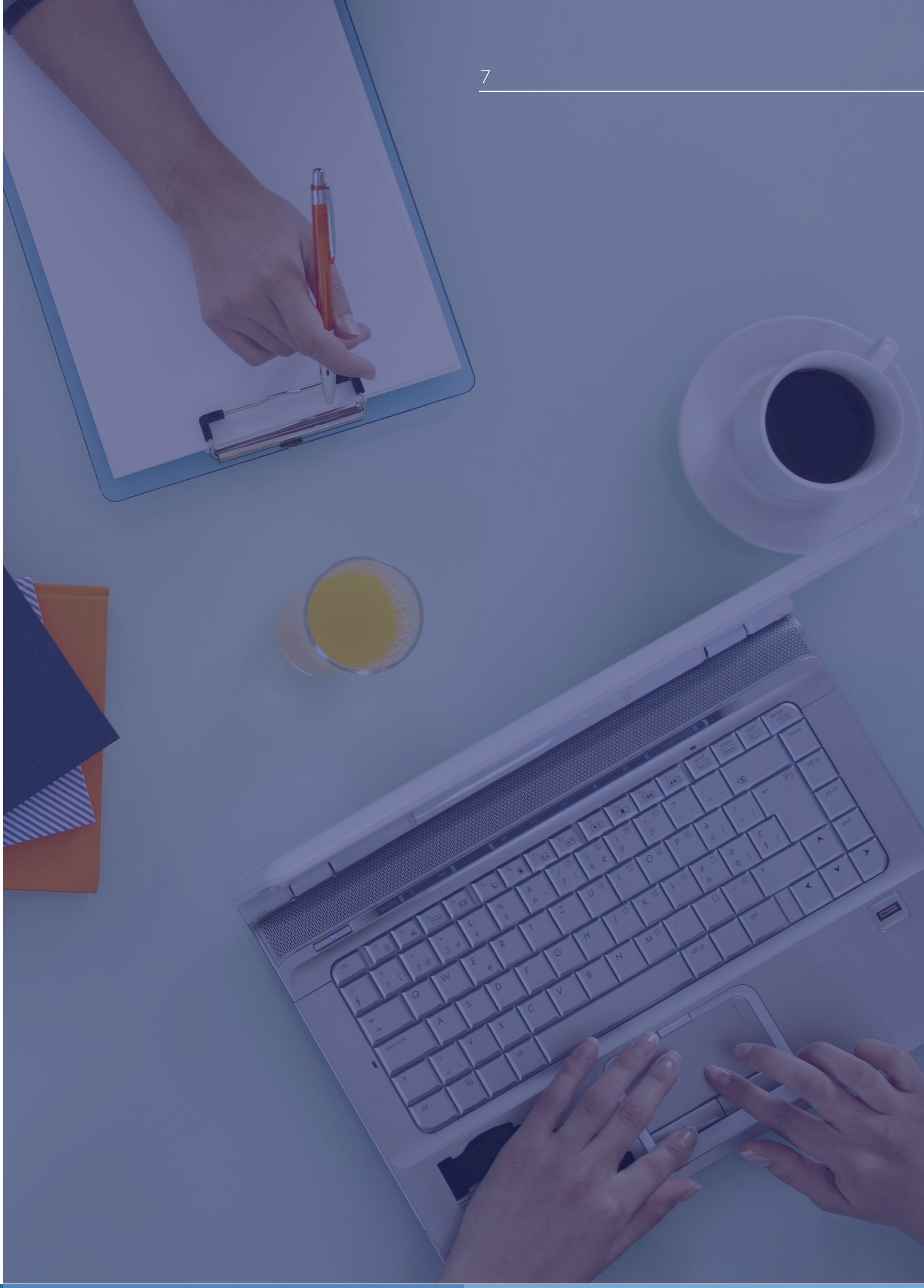
In addition, malware writers are getting sophisticated and reverse engineering some of the latest common threats requires a higher level of knowledge and a larger amount of time dedicated to each sample than historically. Because of this situation antivirus engineers can no longer be employed "by the numbers" to create hundreds of thousands of signatures every few months.

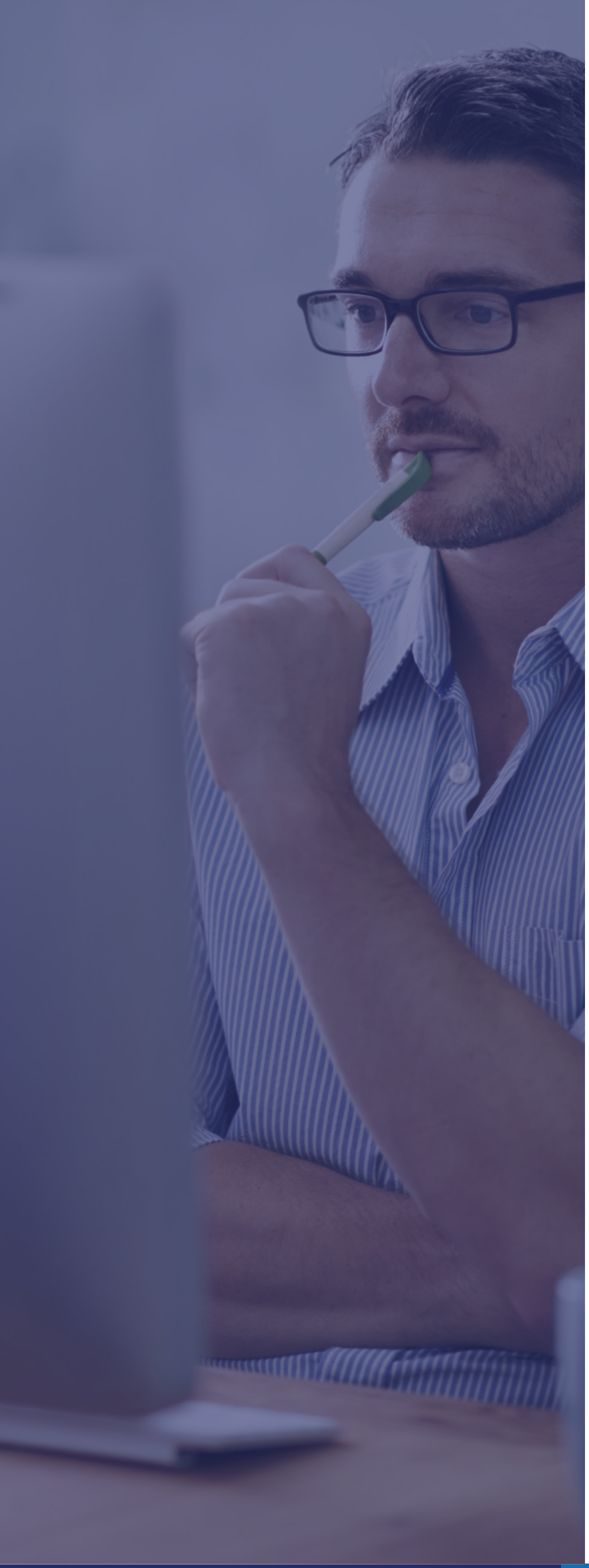
Nowadays, malware only infects a few hundres PCs before updating itself with a new, undetectable variant to avoid detection by regular antivirus signatures. The underlying issue is how does an antivirus lab become aware of such an infection if it is only affecting a handful of users?

Malware techniques and design

The main differences between past viruses and today's malware is that the lifecycle has been significantly shortened and the objectives refined; to steal identities, use computers as spam bots, steal online banking credentials, credit card information, web logins, etc.

More importantly, today's malware is designed to not raise any alarms. Unlike in the past where viruses and worms were designed to spread to as many computers as possible without user intervention. Generating a lot of noise and media awareness whereas, today's criminal malware wants to be as inconspicuous as possible. In order to achieve its objective, malware today uses advanced techniques to evade detection and fly below the radar screen.





1. Targeted Attacks: staying below the radar

One of the main strategies used by Targeted Attacks for staying below the radar is to distribute few copies of many variants[6]. In the past, a single virus or worm was responsible for infecting hundreds of thousands and even millions of computers. Visibility of these situations was very obvious for antivirus labs.

Nowadays, malware only infects a few hundred PCs before updating itself with a new, undetectable variant to avoid detection by regular antivirus signatures. The underlying issue is how does an antivirus lab become aware of such an infection if it is only affecting a handful of users?

2. Malware QA

An older technique used incrementally by malware today is basic QA testing. This is done by testing each variant against the most common antivirus engines to make sure it goes undetected by the majority of them.

This task is greatly simplified by online-scanning services[7] such as Jotti, VirusTotal, the antivirus vendors' online scanning services[7] and online sandboxing services such as Cwsandbox, Norman and Anubis.

Malware creators also count on customized tools to automate testing of new malware against signatures, heuristics and even behavioral analysis technologies. With these tools, malware writers can test the quality of their creations off-line, without risking having the sample sent to the antivirus laboratories via the abovementioned online scanning services.

The objective of malware QA testing is not so much to avoid detection by all scanners and all proactive techniques (generic signatures, heuristics, behavior analysis, behavior blocking, etc.) but to avoid the majority of them. Given its objective of staying below the radar it is not worth creating the most undetected malware if it is only going to live for a few hours or days.

3. Rootkits and sandbox detection techniques

Another common detection evading technique which is gaining momentum[8] is the use of rootkit techniques within Trojan and Spyware samples. When used by malware, rootkits create yet another barrier for being detected, especially as advanced rootkit detection technologies have not yet been deployed to all mass-production security solutions.

It also means that the antivirus laboratories need to spend more time analyzing kernel mode drivers than user-mode samples. For example LinkOptimizer, which has been seen in-the-wild in recent months, is able to determine if the machine it is about to infect has security, debugging or system monitoring tools installed. It also checks if it is running in a Virtual Machine environment. If these checks are matched it silently exits and does nothing.

It also checks if it is running in a Virtual Machine environment. If these checks are matched, it silently exits and does nothing. Labs that depend on VM will have to go through great lengths to be able to install certain LinkOptimizer samples in arder to analyze them in depth.

At the time of writing few anti-malware and security suites include some basic form of rootkit detection such as low-level access cross-view against API-level calls, but most have not yet incorporated the more advanced rootkit detection and deactivation techniques found in free, stand-alone anti-rootkit utilities[9].

Overall the use of rootkits by malware creators keeps growing steadily and this has become a problem for antivirus laboratories that approach malware reverse engineering in a traditional manner and need to analyze each sample one by one.

Antivirus labs are note the only one having problems with rootkit. More and more companies are starting to experience the negative effects of rootkits in business, especially when used for corporate espionage[10].

4. Runtime-packers

Perhaps the most common technique to try to evade detection by anti-malware products is the use of obscure runtime packers with anti-debugging and anti-virtualization techniques.

These types of tools can modify and compress an executable file by encrypting and changing its form from its original format. The final result is a modified executable which, when executed, does exactly the same thing as the original code, but from the outside has a completely different form and therefore evades signature-based detection unless either the engine has the specific unpacking algorithm or it is able to unpack it generically. Malware writers caught up to this approach and we are now seeing malware which use modified versions of known packers or create their own runtime packing routine specifically for their malware samples[12]. In order to address this problem, Panda's engineers have created both generic packer detectors and generic unpacking algorithms which can detect unknown packers and try to unpack them.

However, a more effective solution will be to at least flag the newly created runtime packers as suspicious altogether. Some offthe-shelf perimeter solutions already do this by default. Even some host-based security solutions are using this approach by flagging these types of samples as malicious as they become obvious from the different detection names used by the different anti-malware engines[13]. The impact of such an approach to proactive packer detection comes at a cost.

While speaking to other anti-malware vendors during the 2007 International Antivirus Testing Workshop in Iceland it beca me apparent that doing so in corporate environments was a good approach, but vendors with high installed base, on the consumer market could face such a high wave of false positives that the solution could potentially be worse than the problem itself.

5. Botnets

Botnets are one of today's top threats, causing a significant amount of current infections. For example, over 90% of spam originates from bot-infected computers.

The control of these large networks of compromised computers is sold or rented to perform certain types of cyber-criminal activities: sending spam messages, launching distributed denial of service attacks, renting proxies, stealing user details, etc. In 2010, PandaLabs took part in the shutdown of one of the largest botnets ever reported. This botnet, known as Mariposa, controlled millions of computers in 190 countries, affecting companies, public institutions and home users. The infection was so serious that half of the companies belonging to the Fortune 1000 were compromised by the botnet.

Even though traditional botnets are controlled via IRC, these have evolved a great deal over the last few years. Present botnets are controlled through P2P, HTIP, and even social networking sites like Twitter.

6. Staged infection vectors

It's nothing new that most of today's malware has a tendency of using a two-staged attack as its main infection technique, either by exploiting known or zero-day vulnerabilities or by using small downloaders which change very rapidly to avoid detection.

While in the past it would take malware authors weeks or even months to take advantage of a vulnerability as its main infection vector, nowadays its normal to see exploits in the wild for vulnerabilities a couple of days after it is known. Even further, organizations that manage darknets such as Team Cymru are seeing new zero-day exploits in the wild using stealthier techniques for days and weeks before it is widely known and before they are massively used by botnets.

Examples such as GDI, animated cursor and VML vulnerabilities are being exploited by automated infection frameworks such as Web-Attacker[16] , MPack[17] and Icepack[21] which make use of multiple vulnerabilities to exploit unsuspecting and un-patched users in order to infect them with a Trojan.

Downloaders have also become common practice for two-staged infection techniques. First a small file is executed either via a browser drive-by download or similar exploit. This file is coded with a single objective in mind; download a second file from a URL and execute it. This second file in turn is the true Trojan which ends up infecting the system.

These downloaders have become very advanced. SecuriTeam recently ran a Code Cruncher competition to create the smallest downloader in the world[18]. More recently we are se eing a myriad of graphical tools emerge that simplify the creation of new downloaders[19], even with custom packing techniques to evade detection.

7. "Malware 2.0"

A current trend in malware creatien is that the actual binary that infects the user's PC is "dumb" and the intelligence is "in the cloud". The code that resides en the PC has some simple functions that it passes on to a remotely compromised server. The server then returns instructions on what to do. Borrowing the (perhaps overused) "2.0" term from current web trends, we will refer to "Malware 2.0" as malware which separates its intelligence from its code base.

Pandalabs has reported the "2.0" approach in banking targeted attack Trojans in order to remotely monitor users browsing habits and, based on the online banking landing page and authentication scheme, inject some type of HTML cede or ether. Known banking Troja as such as LimboINetHell and SinowallTerpig use these techniques quite extensively[20].

Other "2.0" techniques used by malware are "server-side-compilation ", where the webserver re-compiles a new binary every few hours. Lastly, botnets are using fast-flux DNS networks for improved resistance against take down efforts. An example of this is the technique used in the StormINuwar attacks.

03 Panda's Technology Evolution

At Panda Security we research and develop 100% of our core anti-malware technologies.

This dedication to innovation has allowed us to lead the way in proactive technology deployment to the market.

Dealing with this malware situation using a traditional signature approach has not been valid for some years now.

A complete Host Intrusion Prevention System (HIPS) with advanced heuristics, deep packet inspection firewall, behavior blocking, behavior analysis and system and application hardening are an absolute must for any security solution.

The sad reality however is that about half the solutions on the market do not have these types of technologies yet[21].

Following a defense-in-depth philosophy, which could be summarized as integrating different protection technology layers at different infrastructure layers, Panda Research, a team dedicated to developing new security technologies, developed a new focus to security protection which is based on the concept of Collective Intelligence.

Our Collective Intelligence technology is designed to complement Panda's integrated desktop, server and gateway protection to take the battle against today's malware dynamic head on and provide the final complement of Panda's ideal protection model.

Before we dive into explaining Collective Intelligence. Let's walk-through the different technology generations on top of which Collective Intelligence is built.

First Generation: Antivirus

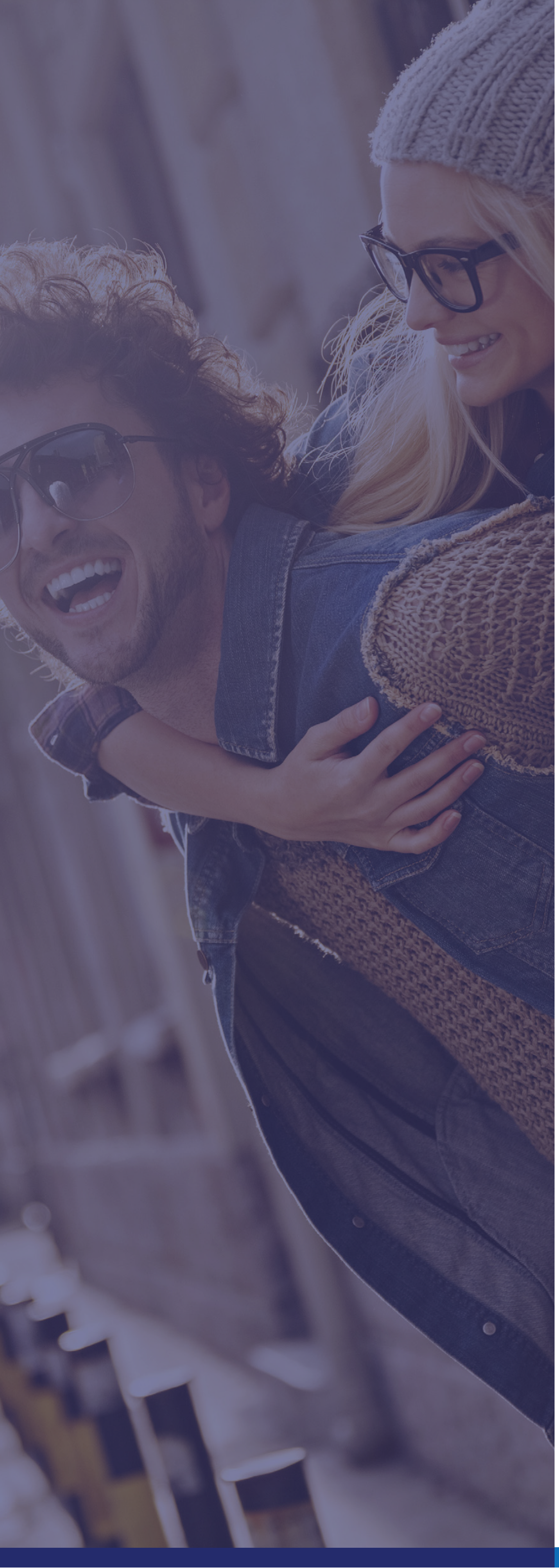
The first generation of antivirus products was purely based on signature detection. This generation of technology occupied most of the 1990's and included polymorphic engines as well as basic rule-based MS-DOS, Win32, Macro and, later on, script heuristics. This period was also marked by the appearance of the first massively used win32 Trojans, such as NetBus and BackOrifice.

Second Generation: Anti-malware

Starting in 2000 new types of malware started to emerge, with file-less network worms and spyware taking the spotlight causing massive and highly visible epidemics.

Basic antivirus engines evolved to integrate personal firewalls to be able to identify and stop network worms based on packet signatures as well as system cleaners to restare modified Operating System settings such as registry entries, HOST files, Browser Helper Objects, etc.

It is within this second generation of technologies that Panda Security integrated the SmartClean functionality into the anti-malware engine, designed to disinfect and restare the Operating System from a spyware or Trojan backdoor infection.



Third Generation: Proactive technologies

Panda released TruPrevent® behavioral technologies in 2004 after more than three years of intensive research and development.

Since then, TruPrevent® has evolved into a set of behavioral technologies that are substantially more effective at blocking zero-day malware proactively without any dependency on viral signatures than any other previous effort in such direction. TruPrevent® is constantly adapted to new malware techniques and exploits.

TruPrevent® was designed as an additional protection layer to the anti-malware engine. Currently there are more than 5 million computers running TruPrevent®.

All these com puters also act as highinteraction honeypot nodes which report to Pandalabs any new malware sample that TruPrevent® flags as suspicious and which is not detected by regular antivirus signatures.

TruPrevent's® approach consists of scanning each item or potential threat using different techniques, carrying out indepth complementary inspections at the different layers of the infrastructure.

The approach to TruPrevent® implementations is modular and therefore can be applied both to desktops and servers to become full-blow integrated Host Intrusion Prevention Systems (HIPS).

As an example of its effectiveness, about two thirds of the new malware samples received at Pandalabs from our users are now coming from automated submissions from TruPrevent®.

Technically, TruPrevent® consists of 2 main technologies: behavioral analysis and behavioral blocking, also known as system and application hardening. Before going into each of these let's take a look at the underlying uncloaking layer which makes malware visible to these behavioral technologies.

1. Uncloaking techniques

As malware has evolved so have the techniques used to evade detection and hide from prying eyes. To combat these hiding techniques there is an underlying lay er of uncloaking technologies common to all of Panda's products.

The following techniques are able to inspect any item as deeply as required, even if the item is making use of stealth techniques to remain hidden in the system, and pass the results on to the scanning and monitoring technologies:

- Deep Code Inspection
- Generic Unpacking
- Native File Access
- Rootkit Heuristic

2. TruPrevent

Behavior Analysis Codenamed Proteus, it acts as a true last line of defense against new malware execut in gin the machine that manages to bypass signatures, heuristics and behavior blocking. Proteus intercepts, during runtime, the operations and API calls made by each program and correlates them before allowing the process to run completely. The real-time correlation results in processes being allowed or denied execution based on their behavior alone.

As soon as a process is executed, its operations and API calls are monitored silently by Proteus, gathering information and intelligence about that process's behavior. Proteus exhaustively analyses the behavior and is designed to block the malware as soon as it starts performing malicious actions. If it is determined as suspicious, the process is blocked and killed before it can carry out all of its actions and prevented from running again.

Unlike other behavioral technologies, Proteus is autonomous and does not present technical questions to the end user (" Do you want to allow process xyz to inject a thread into explorer. exe or memory address abe?"). If Proteus thinks that a program is malicious it will block it without requiring user intervention. Most users cann ot make informed decisions when it comes to security. Some behavioral products throw nondeterministic opinions -or behavioral indecisions- whose effectiveness depends on the user clicking on the right choice. A key functionality of any behavioral technology must be making decisions without user intervention. Anything less is a potential point of failure. Our internal statistics show that this technology alone is capable of detecting over 80% of the malware in the wild without signatures and without generating false positives.

This technology does not require signature updates, as it is based solely on the behavior of applications. A bot would not be a bot if it didn't behave as such, but if it does it will be detected by this technology, regardless of its shape or name. Several third-party tests have been performed on TruPrevent®. Performing tests for behavioral technologies such as TruPrevent, using real-lite malware samples, is timeconsuming and it requires a fair amount of expertise in the field. It is without doubt much more challenging than performing on-demand tests of antivirus scanners against a collection of viruses.

The first test was commissioned by Panda and was performed by ICSALabs, a Division of CyberTrust Corporation, in the fall of 2004. ICSALabs tested the technologies against a set of approximately 100 real malware samples. This first test was designed to verify that the technologies worked against a variety of malware types, rather than to reach a conclusion about the overall effectiveness of the technologies over time. Time, however, has shown that the innovation brought about by TruPrevent Technologies marked the path to follow as demonstrated by independent reviews and analyses carried out over the last few years.

These are just some recent examples:

- In May 201 O, AV-Comparat ives con ducted a "retrospective test" consisting of disabling the antivirus updates and Internet access and attempting to detect malware that appeared a month later than the last update. In such circumstances, an antivirus product can only rely on its technologies and proactive signatures to detect malware. Panda Antivirus Pro was the winner of this comparative review (together with TrustPort) with a 61% detection ratio, 20% higher than the average result[28].

- Also in May 201 O, the German publication c't magazine carried out a "0-day" detection test. In this case, the antivirus programs were exposed to recent malware strains that could not be detected reactivel y by security companies. Panda Cloud Antivirus took top honors in the review, with a 99.10%detection ratio.

3. TruPrevent® Behavior Blocking

Codenamed KRE (Kernel Rules Engine), this is TruPrevent's second main component, also known as Application Control & System Hardening or Resource Shielding.

Hackers and malware abuse the privileges of legitimate applications to attack systems by injecting code. To prevent these types of attacks generically it is very cost-effective to use rule-based blocking technology which can restrict the actions that authorized applications can perform in the system. KRE is composed of a set of policies which are defined by a set of rules describing allowed and denied actions for a particular application of group thereof. Rules can be set to control an application's access to files, user accounts, registry, COM objects, Windows services and network resources. Despite offering a high degree of granularity to administrators for creating custom policies, the Application Control & System Hardening module (KRE) is shipped with a set of default configuration policies which are managed and updated by Pandalabs.

The default policies provide protection against attacks exploiting common weaknesses found in out-of-the-box as well as fully-patched installations of Windows operating systems. A recent example of the effectiveness that proactive blocking provides can be seen with the never-ending wave of PDF format vulnerabilities, affecting Acrobat Reader specially. These flaws have been exploited recently to spread malware and carry out targeted attacks on certain companies.

The detection ratio of files exploiting these vulnerabilities using traditional detection means like signatures is very low. In most cases, it doesn't even reach 50%. On the other hand, behavioral blocking technologies such as TruP reve nt, proactively prevents Microsoft Word, PowerPoint, Excel, Access, Acrobat Reader, Wi ndows Media Player and other applications from dropping and running any type of executable code on the system. Unlike any antivirus signatures tested, Tru Prevent® provi des real zero-day protection against any Microsoft Office exploit, known or unknown.

4. Genetic Heuristics

"Genetic" technologies are inspired by the field of genetics in biology and its usefulness to understand how organisms are individually identified and associated to other organisms. These technologies are based on the processing and interpretation of II digital genes11, which are represented in our case by quite a few hundred characteristics of each file that is scanned.

Codenamed Nereus, the Genetic Heuristic Engine was initially released in 2005. The objective of GHE is to correlate the genetic traits of files by using a proprietary algorithm. The genetic traits define the potential of the software to carry out malicious or harm less actions when executed on a computer. GHE is capable of determining whether a file is innocuous, worm, spyware, Trojan, vi rus, etc. by correlating the different traits of each item scanned. GHE can be set to low, medium or high sensitivity with the obvious combination trade-off between detection rates and false positives. The different sensitivity levels are designed to be applied to different environments depending on the probability of malware prevalence each.

For example, at network SMTP gateways we have found that the likelihood of an executable file being malware is very high. Therefore, the implementation in our commercial products is of high sensitivity for network layer e-mail scanning products. However for storage (or applications) layers where the vast majority of executable code is from legitimate applications, we have implemented GHE with medium sensitivity. With this setting we've been a ble to maximize detection rates for unknown malware while resulting in a negligible false positive rate.

The results of GHE so far are excellent. Since its release, rough ly one third (cumulative) of the new variants received at Pandalabs from real users' machines have been submitted automatically by the GHE.

Fourth Generation: Collective Intelligence

As we have often said in the past, the amount of malware in circulati on is increasing significantly year after year. Unlike other metrics, which tend to show a "war of numbers" between security companies inst ead of objective information, we all agree on the fact that every year we receive and detect as much malware as in all previous years combined.

This means that a security solution today must be able to detect forty times more malware than five years ago, or three times more than just two years ago.

While a full-fledged HIPS solution raises the bar substantially by detecting and blocking most of th is ma lwa re with proactive technologies, it is still possible for unknown malware to slip through its defenses. We need to consider the fact that, while 80% or 90 % of proactive effectiveness is relatively speaking an excellent score, in absolute terms it may lead to hundreds or thousands of malware samples being missed in 201 O, since even a small fraction of a large enough number will still be a 'big' number.

Also, you must bear in mind that the huge amount of new malware in circulation doesn't necessarily mean there is a huge presence of active malware at a given moment. The strategy followed by malware creators has changed. Generally, they are no longer interested in creating malware that triggers massive infections and last indefinitely, but malicious code that spreads very quickly and in waves. That is, the malware variants they create are quickly replaced by others sometimes even before traditional, signa ture-based technologies can detect them.

As a consequence. the average lifespan of malware has dropped from weeks or months to days ar even hours. Under these circumstances, a protection approach based on signature file updates published every 24 hours would leave users unprotected during most of the time that each malware strain is active.

Panda's response to the challenges posed by the new malware dynamic (increasingly large volumes of malware combined with limited distribution and extremely short lifespans) continues to be Collective Intelligence.

1. Functionalities

Panda's Collective Intelligence technology was initially released at the end of 2006 in limited pilots with the objective of being a ble to reliably detect "10 times more than we were detecting with 10 times less effort". Collective Intelligence functions as an online and real- time Security-as-a-Service (SaaS) platform. After two years of preliminary research and development and another two years of exploitation and evolution of the marketed product, the technology has produced excellent results with regards to malware received, analyzed and detected, as wel l as anti-malware response time.

Our Collective Intelligence technology integrates different functional modules providing complementary services.

- Real-time sensor network.
- Automated malware collection.
- Automated malware processing and classification.
- Automated malware remediation.

Below is a description of each of these modules. It is not our intention to give a technical overview of the system architecture, but a high-level explanation of their functionalities.

a. Real-time sensor network

Thanks to the presence of advanced sensors in our products, the Collective Intelligence agent collects information about memory processes and objects and sends queries to the CI central servers, which, in turn, store and correlate the information received in real time from users' computers. This gives us great visibility into the appearance and behavior patterns of new executable files on the computers of our community of users, making it possible to identify potentially suspicious processes.

As soon as a file is identified as suspicious, the co rresponding information is immediately made available to the community. This prevents having to carry out similar scans on other PCs and stops malw are from sp rea ding to other computers before being detected.

b. Automated malware collection

If certain conditions are met, the suspicious files or parts of then are automatically uploaded to the CI servers where they are processed further.

Since processes loaded in memory are not subject to many of the cloaking techniques and "reveal" themselves, the agent component does not need to contain a large amount of intelligence and uncloaking routines and can therefore be very light. Additionally, virtually all companies in the security industry as well as top analysts regularly exchange malware samples in order to keep malware at bay with everyone's joint effort. The CI malware collection systems let Panda automate these data exchanges , improving processing time and the quality of the information received and sent to other security vendors.

c. Automated malware processing and classification

Cloud-based processing is not limited by the CPU and memory constraints of personal computers. Therefore, scanning routines at the CI server undergo much more in-depth processing by more sensitive technologies (signature and sensitive heuristics scanning, emulation, sandboxing, virtualization, whitelisting, etc.) to reach a final classification. It is important to note that the scanning power used at the CI servers is only limited by hardware and bandwidth scaling, unlike a typical scenario on a PC, desktop computer or server. Therefore many of the more res o urce-in ten sive p reactive techniques which Pandalabs is using, and which provide much higher detection rates (at an also higher computational costs) can now be used massively for the benefit of the users without even touching valuable customer's CPU and memory resources ..

With this approach, the majority of new malware samples can be analyzed and classified automatically in a matter of minutes. The CI servers are managed by Pandalabs allowing samples that cannot be classified automatically to the ultimately looked at by an analyst at the lab.

d. Automated malware remediation

The CI remediation module is in charge of automatically creating detection and disinfection signatures for the samples previously analyzed by the processing and classification module. These signatures are in turn used by the community of CI users to proactively detect and disinfect new or even targeted attacks with very low numbers of infected hosts.

Traditional HIPS and anti-malware solutions also benefit from Collective Intelligence. The remediation module has generated hundreds of thousands of ma lw are signatures that we have gradually deployed into our existing products.

2. Benefits

e. Community leverage

Traditional security solutions are architected with a PC-centric philosophy. This means that a PC is treated as a single unit in time and any malware detected within that PC is considered separately from the rest of the malware samples detected in millions of other PCs.

Traditional security companies do not have visibility into what PC a particular piece of malware was first seen on. Neither is there visibility of the continuity of that malware's evolution over time on different PCs. Most importantly, other PCs do not automatically benefit from proactive malware detections on them. Thanks to the CI technology, all evidence of malware presence on one PC is immediately leveraged by the others. This way, the Panda user community works as an early warning system, producing a much larger knowledge base than the one obtained through the individual, isolated scanning of each PC.

f. Increased malware processing capacity at our labs

One of the biggest barriers to raising the bar of reliable malware detection ratios was the fact that the process of creating a signature against a single sample took too long. Each malware sample needed to be sent to the lab by an affected user or fellow researcher, and reversed engineered by a lab technician which in turn needed to create a detection signature and disinfection routine for it. Up until the implementation of the automatic malware processi ng, classification and remediation modules, the entire process was, in most cases, mostly manual and could take anywhere from minutes, to hours or days or even weeks, depending on the lab engineers workload and other factors such as sample priority, preva lence, da mage potential, media coverage, etc.

It was a cumbersome and time consuming process that had to be carried out for each malware strain. Multiply that effort by thousands of new samples every day, and you will realize that it would be impossible to stop the current malware avalanche using such a methodology.

The Collective Intelligence infrastructure allows the entire process to be automated and performed within seconds for most samples, allo wing lab engineers to concentrate on analyzing particularly complex specimens or generate advanced, more proactive and general detection routines.

g. Reduced bandwidth and disk space use

One of the main benefits of Col lective Intelligence is that it eliminates the need to download signatures to each customer as they are available in the cloud.

The number of det ection routines generated by Collective Intelligence since it was first implemented is so huge that if you wanted to put all that knowledge into traditional signature files you would need 250 MB. This would make it impossible to handle daily file updates smoothly.

Thanks to the knowledge available on our servers we can keep significantly smaller signature files that nevertheless include enough information about currently active malware as to provide effective protection even if a user's computer is not connected to the Internet at a given time.

h. Increased reaction speed to newmalware

As previously explained, the traditional approach involved analyzing each PC as an isolated entity. This not only resulted in less global knowledge of the malware in circulation, but increased the time that lapsed between the appearance of a new malware strain and the availability of the relevant detection routine. According to this approach, users had to wait for the antivirus lab to receive that specific sample, and for a signature to be created, QA'ed and deployed in order to be protected. Ultimately, this resulted in traditional approaches being too slow to combat today's rapidly evolving malware.

One of the main advantages of Collective Intelligence besides the automation of the malware removal cycle, is the benefits provided to the user community. As soon as the Collective Intelligence servers determine that an executable file is malicious, the relevant knowledge is made available to the entire user base. The protection is deployed to all users within minutes in comparison to the 24-hour cycles typical of most traditional approaches.

i. Gaining knowledge on malware techniques

Another main benefit provided by the comm unity fe ature of Col lective Intelligence is of giving our engineers insight into new malware techniques and entry points. Questions such as where a specific piece of malware was first found and how it spread allow us to model additional intelligence into specific malware families and even creators of specific malware variants. This approach of applying data warehousing and data mining techniques to malware detections by the community provides significant knowledge on how malware and targeted attacks are carried out. The type of knowledge that can be gathered using this approach becomes especially useful if it can be applied for tracking infection origins, which in turn might have some interesting applications and benefits for law enforcement efforts.

3. A few significant facts about Collective Intelligence

The system receives over 17 million information requests every day from Panda solution users.

It processes over 75.000 new, unknown files every day and determines whether they are malware or not.

An average of 55.000 new malware strains appear every day. The total number of catalogued malware samples exceeds 40.000.000.

0.6% of malware strains that are not automatically categorized by the system are manually analyzed by Panda technicians.

At present, the Collective Intelligence main database occupies 2.5 TB of disk space (considering categorized malware only) and generates 190 GB of logs every day.

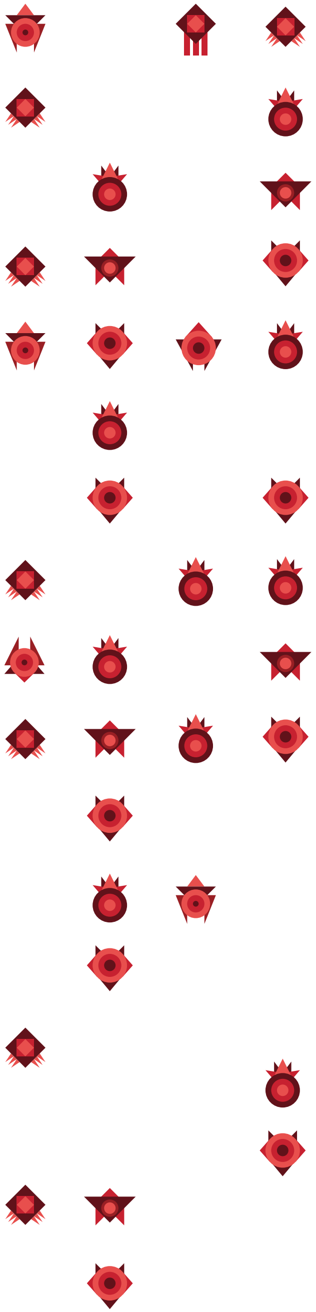
The Collective Intelligence database contains over 1.000.000.000.000 records.

4. Deploying Security Services "from-the-cloud"

We have developed and deployed a series of security services that function purely based on the Collective Intelligence platform. These online services are designed to perform in-depth audits of machines and detect malware not detected by the installed security solution.

For home users we have deployed Panda Cloud Antivirus Pro (www.cloudantivirus.com), which scans of the entire computer including hard drive, memory, email databases, etc.

On the business front, the requirements for performing and in-depth malware audit are more demanding. Therefore we have created a specific managed service called Malware Radar (www.malwareradar.com). Thanks to this service companies can quickly perform complete audits of their entire network endpoints to verify their level of security, pinpoint non-detected infection sources or to unveil machines which have been subject to targeted attacks.



04 Conclusion

The latest advances by the black hat and cybercrime communities are taking advantage of the inherent weaknesses in the security industry:

- The labs are being swamped by more malware which is being created every day.
- By remaining invisible users do not perceive the need for additional protection.
- Targeted attacks that only infect very few users are more effective than epidemic attacks that infect millions of users.
- Users tend to trust a single solution or single layer of protection as their main line of defense against malware.

As malware techniques advance in this cat-and-mouse game, security vendors need to add more layers of protection to keep customers safe. The need for additional protection is revealed by the fact that a large portian of users with current and updated security solutions is in fact infected.

To tackle today's problem we need new layers of protection that take advantage of automating the entire malware protection cycle, from sample collection, analysis, classification to remediation. But automation by itself is not enough. We also need visibility into what's happening on all PCs in arder to detect targeted attacks more efficiently and gain a competitive edge on malware creators.

The technology developed by Panda Security, called Collective Intelligence, provides all the benefits of an added layer of defense that provides effective response and protection to the current malware threats, is able to detect targeted attacks and gains intelligence thanks to the correlation of all the detections by the community of users.

05 References

1. Research Study: Active Infections in Systems Protected by Updated AntiMalware Solutions. Panda Research. http://research.pandasecurity.com

2. Gartner's 10 Key Predictions for 2007. Gartner. http://www.eweek.com/article2/0,1895,2072416,00.asp

3. The Zero-Day Dilemma. Security IT Hub. http://www.security.ithub.com/article/The+ZeroDay+Dilem ma/199418_ 1.aspx

4. Welcome to 2007: the year of professional organized malware development. F-Prot's Michael St. Neitzel at Hispasec. http://blog.hispasec.com/vi rustotal/16

5. Call the cops: We're not winning against cybercriminals. ComputerWorld. http://www.computerworld.com/action/article.do?command=viewArticleBasic&articleId=9010041

6. The Long Tail: malware's business model. Panda Research. http://research.pandasoftware. com/blogs/research/archive/2 007/01/08ffhe-Long-Tail_3AOO_-malware_2700_s-busi ness model.aspx

7. List of online scanners. CastleCops Wiki. http://wiki.castlecops.com/Online_antivirus_scans

8. Kernel Malware. F-Secure. http://www.f-secure.com/weblog/archives/archive-022007.html#00001118

9. Anti-rootkit.com. List of Rootkit Detection & Removal Software. http://www.antirootkit.com/software/index. htm

10. Rootkit used in Vodafone Phone Tapping Affair. http://www.antirootkit.com/blog/2007/07/12/rootkit-used- in-vodafone-phone-tappi ng-affair/

11. Panda Anti-Rootkit. http://research.pandasoftware. com/blogs/research/archive/2 007/04/27/New-Panda-Anti 2DOO Rootkit- 2DOO -Version- 1.07.aspx

12. Packing a punch. Panda Research. http://research.pandasoftware. com/blogs/research/archive/2 007/02/12/Packing-a-punch.aspx

13. AV performance statistics. OITC & MIRT. Real-time feed of antivirus zero-day detection. http://winnow.oitc. com/avcentral. html

14. Attack of the Zombie Computers Is Growing Threat. The New York Times. http://www.nytimes.com/2007ID 1/07ltechnology/07net. ht ml?ex=1325826000&en=cd 1e2d4c0cd20448&ei=5090

15. 30 Days of Bots Inside the Perimeter. Support Intelligence. http://blog.support-intelligence.com

16. Web-Attacker Exposed. Websense. http:/M11MN.1111ebsense.com/securitylabslblogA:llog.ph p?Blog1D=94

17. MPack Uncovered. http://blogs.pandasoftware. com/blogs/images/Pandalabs/2 007/05/11/MPack.pdf

18. The World's Smallest Downloader. Symantec. http://www.symantec.com/enterprise/secu rity_response/we blog/2006/12/worlds...,smal lest_downloader .html

19. Packing a punch (11). Panda Research. http://research.pandasoftware. com/blogs/research/archive/2 007/03/20/Packing-a-Punch-_2800J IL2900_.aspx

20. Banking Targeted Attack Techniques. Panda Research. http://research.fJandasoftware. com/blogs/images/Panda eCrime2007.pdf

21. Host-Based Intrusion Prevention Systems (HIPS) Update: Why Antivirus and Personal Firewall Technologies Aren't Enough. Gartner. http://www.gartner.com/teleconferences/attributes/attr _ 165 281_115.pdf

24. The Last Great Security Crisis. http://www.eweek.com/article2/0,1895,2095118,00.asp

25. Comments on "The Decline of Antivirus and the Rise of White-Listing". The Register. http://www.theregister.co.uk/2007/06/27/whitelisting_v_an tivirus/comments/

26. "More on White-listing". Kurt Wismer. http://anti-virus rants.blogspot.corn/2007/06/more-on wh1telisting.html

27. http://pandalabs.pandasecurity.com/blogs/iamges/Pandalabs/200712/18/Icepack:. pdf

28. http://www.av comparatives.org/images/stories/tesVondreVavc_report26.pdf

panda

pandasecurity.com