# Combating Malware: Leveraging the Power of a Planet

PANDA SECURITY | *One step ahead.*

# Combating Malware: Leveraging the Power of a Planet

## *Executive Summary*

The amount of malicious software (or malware) that is being released in the world is increasing at an alarming rate. To combat this threat, most antivirus and anti-malware solution vendors are relying on the creation of signatures to protect users. But creating signatures can be a time-intensive process. As a result, current solutions are proving to be much less effective against the proliferation of threats in circulation. Even users protected by solutions with the latest signature databases are frequently infected by active malware. Complementary approaches and technologies must be developed and implemented in order to raise the effectiveness of these solutions to adequate levels.

Collective Intelligence offers a radically different approach to security. This approach is based on exhaustive remote, centralized, and real-time knowledge about malware and non-malicious applications maintained through the automatic processing of all scanned elements. The Collective Intelligence approach provides the ability to maximize malware detection capabilities, while at the same time, minimizing the resource and bandwidth consumption of protected systems.

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 of itself is not enough to tackle the malware cat-and-mouse game. In addition to the large volumes of malware, enterprises also face targeted attacks. Response time in these scenarios cannot be handled by automation of signature files alone. Collective Intelligence provides visibility and knowledge into the processes running on all of the computers scanned. This broad visibility of the community – in addition to automation – is what delivers the ability to tackle not only the large volumes of new malware, but also targeted attacks.

This paper examines the current malware landscape and reviews current emerging malware techniques and design. Topics include the first generation of security solutions – antivirus; the second generation- anti-malware; the third generation – proactive technologies; and finally the fourth generation – collective intelligence solutions.

# Table of Contents

# The Malware Landscape

Security professionals all know that there are more malware samples infecting users than ever before. Malware writers have become more sophisticated and have realized they can obtain large amounts of money from distributing malicious software. The shift in motivation for creating malware, combined with the use of advanced scripting techniques, has resulted in an exponential growth of criminally professional malware being created for the sole purpose of infecting unsuspecting users.

This new malware dynamic has become the next big plague for users and companies alike. Gartner estimates that by the end of 2007, 75% of enterprises will be infected with undetected, financially motivated, targeted malware that has evaded their traditional perimeter and host defenses[1].

In July of 2007, Panda Security conducted a research study[2] to analyze the effectiveness of current anti-malware solutions. The two-part study scanned the PCs of 1.4 million users in over 80 countries. These PCs were utilizing security solutions from over 40 different vendors. Results indicated that even though many consumer PCs tested had protection installed with up-to-date signature databases, nearly 23% of the PCs scanned were infected with malware loaded into memory[3]. The study also examined more than 1,200 firms with security solutions installed. Results showed that nearly 72% percent of companies with more than 100 computers had active malware on their networks.

| Network Size | Rate of infected networks |
|---|---|
| Less than 50 workstations | 36.63 % |
| 51 to 100 workstations | 62.26 % |
| **Over 100 workstations** | **71.79 %** |

Figure 1. The percentage of companies where malware infections were found.

## *Antivirus Laboratories Are Under Attack*

Today's antivirus laboratories are under constant and increasingly frequent attacks. Security industry labs are 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 process from a corporate and business perspective.

Some antivirus solution vendors are trying to deal with the proliferation of malware by increasing the number of analysts at their labs. But malware writers are getting more 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 before. Because of this situation, antivirus engineers can no longer be employed simply "by the numbers" to create hundreds of thousands of new signatures every month.

1 http://blog.gartner.com/blog/predicts2007.php
2 http://www.pandasecurity.com/enterprise/downloads/white-papers/?sitepanda=particulares

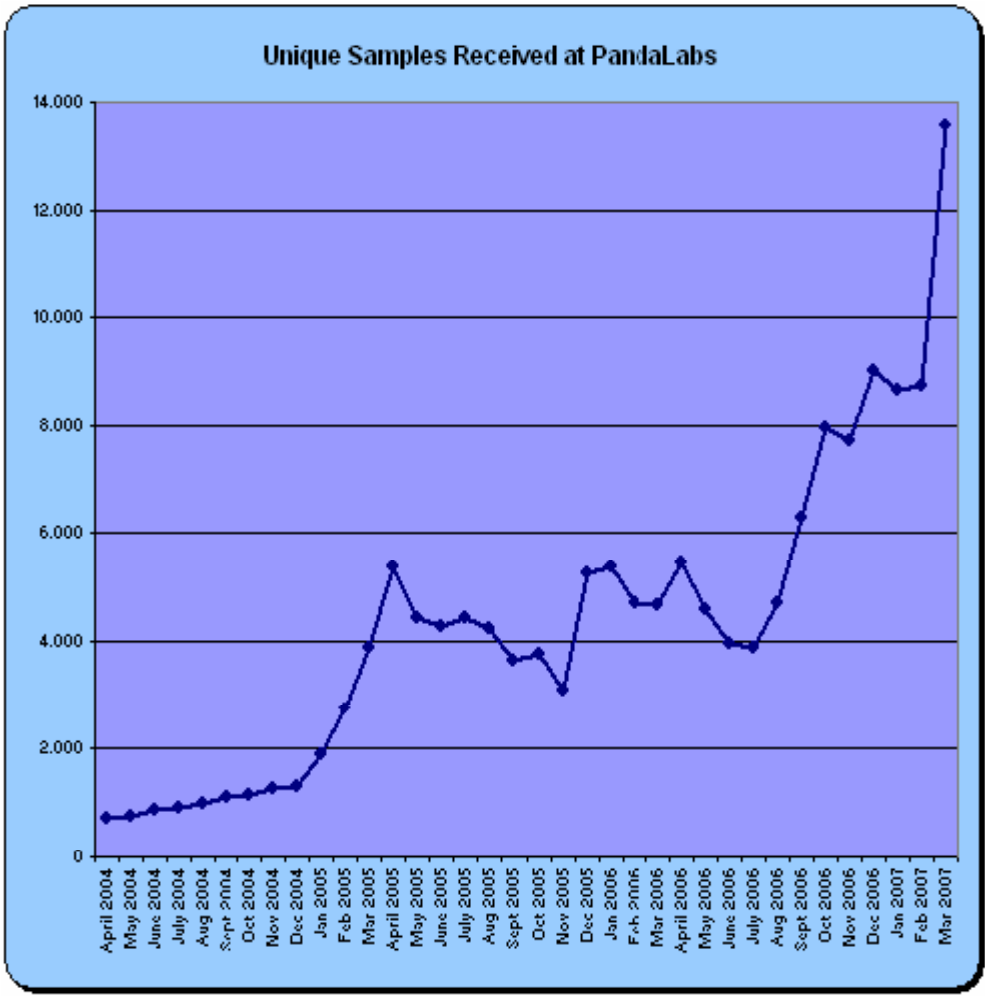


Figure 2. Unique samples received at Panda Security 2004 to 2007.

The security industry is advocating for stronger intervention by law enforcement agencies to convict the most active malware creators. Initiatives to get law enforcement more involved—although definitely a step in the right direction—unfortunately present an insufficient solution, as the number of variants is increasing incrementally and most of the time only the less sophisticated "mules" and "script kiddies" are 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 to battle the proliferation of malware.

## *Malware Techniques and Design*

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

sections will review some of the current approaches to malware, including targeted attacks, malware QA testing, rootkits and sandbox detection techniques, runtime-packers, botnets, staged infection vectors, and malware "2.0" techniques.

## Targeted Attacks: Staying Below the Radar

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 – today's criminal malware is designed to be as inconspicuous as possible.

Malware creators now use advanced techniques to evade detection and to "fly low." One of the main strategies used for staying below the radar is to distribute just a few copies of many variants. 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 the antivirus labs.

Malware designed with this approach may only infect a few hundred PCs before updating itself with a new, undetectable variant to avoid detection by regular antivirus signatures. The underlying challenge becomes "how does an antivirus lab become aware of such an infection if it is only affecting a handful of users?"

## Malware QA Testing

An older technique used incrementally by malware propagators 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 such as Jotti, VirusTotal, the antivirus vendors' online scanning services, and online sandboxing services such as Cwsandbox, Norman, and Anubis.

Today's malware creators also count on customized tools to automate testing of new code 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 above-mentioned online scanning services.

The objective of malware QA testing is not so much to avoid detection by all scanners and all proactive techniques, but to avoid the majority of them. Given the objective of staying below the radar, it is not worth creating the most sophisticated and undetectable malware if it is only going to live for a few hours or days.

## Rootkits and Sandbox Detection Techniques

Another common detection evading technique which is gaining momentum 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 is able to determine if the machine it is about to infect has security, debugging, or system monitoring tools installed. It also checks to see if it is running in a virtual machine (VM) 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 order to analyze malware in depth.

Currently, only a few anti-malware and security suites include some basic form of rootkit detection, such as low-level access cross-view against API-level calls. Most have not yet incorporated more advanced rootkit detection and deactivation techniques found in free, stand-alone anti-rootkit utilities.

Overall the use of rootkits by malware creators is steadily growing. This has become a problem for antivirus laboratories that approach malware reverse engineering in a traditional manner and analyze each sample one by one. Not only are the antivirus labs having problems with rootkits, but also companies are starting to experience the negative effects of rootkits in business, especially when used for corporate espionage.

## Runtime-Packers

One of the most common techniques used 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 does exactly the same thing as the original code when executed, but from the outside has a completely different form. This enables it to evade signature-based detection unless either the engine has the specific unpacking algorithm or is able to unpack it generically. Malware writers have caught up to this approach and are now creating code which uses either modified versions of known packers or creating their own runtime packing routine specifically for their malware samples.

In order to address this problem, 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 is to flag the newly created runtime packers as altogether suspicious. Some off-the-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. But the impact of such an approach to proactive packer detection is not without cost. Vendors with a large 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.

## Botnets

A large portion of money made by cyber-criminals stems from botnets[3]. According to recent studies, approximately 11% of computers worldwide are infected by bots which are responsible for sending up to 80% of all spam. The control of these large networks of compromised machines is sold or rented to perform certain types of cyber-criminal activities from sending spam runs, distributed DoS attacks, renting of proxies, keylogging, pay-per-click installs, adware installations, stored passwords, man in the middle attacks, etc.

This example illustrates how a botnet is created and used to send email spam:

1. A botnet operator sends out viruses or worms, infecting ordinary users' computers, whose payload is a trojan application – the bot.
2. The bot on the infected PC logs into a particular server. That server is known as the command-and-control server (C&C).
3. A spammer purchases access to the botnet from the operator.
4. The spammer sends instructions via the server to the infected PCs

3 Botnet is a jargon term for a collection of software robots, or bots, which run autonomously and automatically. They run on groups of "zombie" computers controlled remotely by crackers. This can also refer to the network of computers using distributed computing software.

5. Causing them to send out spam messages to mail servers.

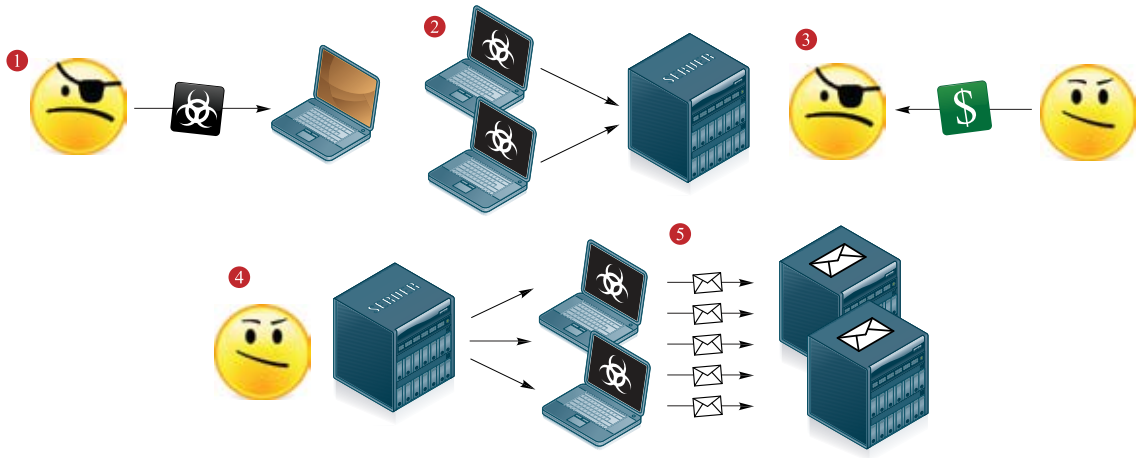


Figure 3. Botnet Formation and Exploitation.

Panda Security has witnessed on-line wars between different bot gangs to win over hijacked PCs. Even though evidence suggests that there are many PCs belonging to Fortune 500 companies that are controlled remotely by bot herders to send out spam, the reality is that virtually every corner of the Internet is infested by bots.

## Staged Infection Vectors

Most of today's malware uses a two-staged attack as its main infection technique, either by exploiting known or zero-day[4] vulnerabilities or by using small downloaders which change rapidly to avoid detection. While in the past it would take malware authors weeks or months to take advantage of a vulnerability as the main infection vector, now it is normal to see exploits in the wild[5] for vulnerabilities a couple of days after they are 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 or weeks before they are widely known and massively used by botnets.

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. Malware creators are now using a myriad of emerging graphical tools that simplify the creation of new downloaders, even with custom packing techniques to evade detection.

## Malware 2.0

A current trend in malware creation is that the actual binary that infects the user's PC is "dumb" and the intelligence is "in-the-cloud" (on the Internet). The code that resides on the PC has some simple functions that it passes on to a remotely compromised server. The server then returns instructions on

4 A zero-day (or zero-hour) attack is a computer threat that exposes undisclosed or unpatched computer application vulnerabilities. Zero-day attacks take advantage of computer security holes for which no solution is currently available.0-day exploits are released before the vendor patch is released to the public.

5 The phrase 'exploits in the wild' refers to viruses or malware that is now outside the single computer or lab where it was created.

what to do. The term "Malware 2.0" refers to malware which separates its intelligence from its code base.

Panda Security has observed the 2.0 approach in banking-targeted attack Trojans in order to remotely monitor users' browsing habits. Then, based on the online banking landing page and authentication scheme, the malware injects some type of HTML code or other. Known banking Trojans (such as Limbo/NetHell and Sinowal/Torpig) use these techniques quite extensively. Other 2.0 techniques include "server-side-compilation" where the Web server recompiles a new binary every few hours. Lastly, botnets are using fast-flux DNS networks for improved resistance against take-down efforts.

# Panda's Technology Evolution

Dealing with malware using a traditional signature-based approach is no longer sufficient. 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. Unfortunately, only about half of the solutions on the market today have these types of technologies. Panda Security's solutions employ all of these techniques.

Panda Security has developed a robust, defense-in-depth philosophy to security protection integrating different protection technology layers at different infrastructure layers, based on the concept of Collective Intelligence[6] (CI). The CI concept complements Panda's integrated desktop, server, and gateway protection to take the battle against today's malware dynamic head on and provide the final complement to Panda's ideal protection model. But before explaining Collective Intelligence in depth, this paper will review the different technology generations on top of which Collective Intelligence is built.

## *First-Generation Security Solutions: Antivirus*

The first generation of antivirus products was purely based on signature detection. This generation of technology occupied most of the 1990s 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 Security Solutions: Anti-Malware*

Starting in 2000, new types of malware started to emerge with fileless network worms and spyware taking the spotlight, causing massive and highly visible epidemics. Basic antivirus engines evolved to integrate personal firewalls to identify and stop network worms based on packet signatures, as well as system cleaners to restore modified operating system settings such as registry entries, HOST files, Browser Helper Objects, etc. It is within this generation of technologies that Panda Security integrated its SmartClean functionality into the anti-malware engine, designed to disinfect and restore the OS from a spyware or Trojan backdoor infection.

## *Third-Generation Security Solutions: Proactive Technologies*

Third-generation security solutions employ a variety of proactive technologies. Panda TruPrevent® is a third-generation solution that provides a set of behavioral technologies that is very effective at blocking zero-day malware proactively without any dependency on viral signatures. More than any other previous efforts in this direction, TruPrevent is constantly adapted to the ever-changing new malware techniques and exploits.

6 Collective intelligence is a form of intelligence that emerges from the collaboration and competition of many individuals.

TruPrevent delivers an additional, highly protective layer to the anti-malware engine. Currently there are more than five million computers running Panda Security's TruPrevent solution. All of these computers act as high-interaction honeypot7 nodes which report back to Panda any new malware samples flagged as suspicious and not detected by regular antivirus signatures.

TruPrevent's approach consists of scanning each item or potential threat using different techniques and carrying out in-depth complementary inspections at the different layers of the infrastructure. The approach to TruPrevent implementations is modular and can be applied to both desktops and servers to become a full-featured, integrated HIPS. As a true measure of its effectiveness, about two thirds of the new malware samples received at Panda Security from users are now coming from automated submissions from TruPrevent.

TruPrevent consists of two main technologies: behavioral analysis and behavioral blocking, also known as system and application hardening. Before going into each of these in detail, this paper will take a look at the underlying uncloaking layer which makes malware visible to these behavioral technologies.

## Uncloaking Techniques

As malware has evolved, so have the techniques that are used to evade detection and hide from prying eyes. To combat these hiding techniques, there is an underlying layer of uncloaking technologies common to all of Panda products. Several techniques are able to inspect any item as deeply as necessary – even if the item is making use of stealth techniques to remain hidden in the system – and pass on the results to the scanning and monitoring technologies. These techniques include deep code inspection, generic unpacking, native file access, and rootkit heuristics.

## Behavior Analysis

The TruPrevent Behavior Analysis module acts as a last line of defense against new malware executing in the machine that manages to bypass signatures, heuristics, and behavior blocking. During runtime, the module intercepts the operations and API calls made by each program and correlates them before allowing the process to run completely. This 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 the Behavior Analysis module, gathering information and intelligence about that process's behavior. The module exhaustively analyses the behavior and blocks the malware as soon as it starts performing malicious actions. If it is suspicious, the process is blocked and killed before it can carry out its actions and is prevented from running again.

Unlike other behavioral technologies, the Behavior Analysis module is autonomous and does not present technical questions to the end user (e.g., "Do you want to allow process xyz to inject a thread into explorer.exe or memory address abc?") If the module thinks that a program is malicious, it will block it without requiring user intervention. Studies show that most users cannot make informed decisions quickly when it comes to security. Some behavioral products include non-deterministic opinions (or behavioral indecisions) whose effectiveness depends on the user clicking on the right choice. To be effective, key functionality of any behavioral technology must be making decisions without user intervention. Anything less introduces a potential point of failure.

7 Honeypots are servers or devices that expose server services that wait passively to be attacked in search of malicious servers that attack clients.

Panda Security's internal statistics show that this technology alone is capable of detecting over 80 percent of the malware in the wild – without creating 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 so, it will be detected by this technology regardless of its shape or name.

Several third-party tests have been performed on TruPrevent. The first test was commissioned by Panda and was performed by ICSALabs, a Division of the CyberTrust Corporation. ICSALabs tested the technologies against a set of 100 real malware samples. The 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. At the same time, ICSALabs tested TruPrevent against several sets of legitimate applications, from games to peer-to-peer packages, but was not able to produce any instance of false positives. Another "early" review by PC Magazine USA concluded that "TruPrevent blocked two-thirds of a sample of recent worms, viruses, and Trojans based strictly on behavior. Blocked no legitimate programs. No noticeable impact on system performance."

The following diagram illustrates the path that network or Internet traffic would follow in a system protected by Panda Security's TruPrevent integrated endpoint security technology.

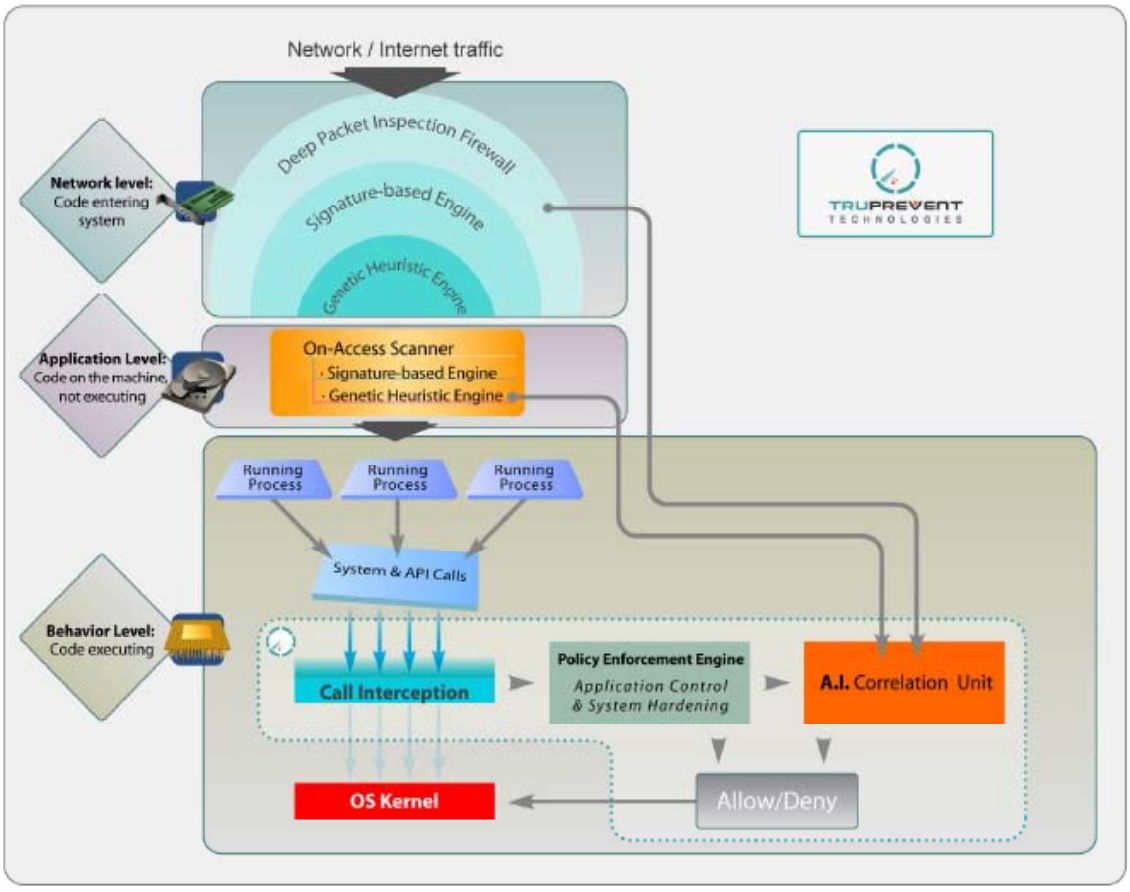


Figure 4. Panda Security's TruPrevent integrated endpoint security technology.

## Behavior Blocking

The Kernel Rules Engine (KRE) is TruPrevent's second main component, also known as "Application Control and 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, it is very cost-effective to use generic 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 rules describing allowed and denied actions for a particular application or group thereof. Rules can be set to control an application's access to files, user accounts, registry, COM objects, Windows services, and network resources.

In addition to offering a high degree of granularity to administrators for creating custom policies, KRE provides a set of default configuration policies which are managed and updated frequently by Panda Security. 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 is the never-ending wave of Microsoft Office format vulnerabilities which are being exploited to hide malicious code. These vulnerabilities are being used by skilled, targeted attacks. According to a study of known (patched) and zero-day (un-patched) Microsoft Office vulnerability exploits, an average antivirus signature detection rate of only 50% was achieved by a variety of tested antivirus engines.

However, behavioral blocking technologies (such as TruPrevent) proactively prevent from the onset all MS Word, PowerPoint, Excel, Access, Acrobat Reader, Windows Media Player, and other applications from dropping and running any type of executable code on the system. Unlike any antivirus signatures tested, TruPrevent provides real zero-day protection against any Microsoft Office exploits – both known and unknown.

## Genetic Heuristics

Genetic heuristic technologies are inspired by the field of genetics and its ability to reveal how organisms are individually identified and associated to other organisms. These technologies are based on the processing and interpretation of "digital genes" which are represented in this case by quite a few hundred characteristics of each file that is scanned.

Panda Security's Genetic Heuristic Engine (GHE) correlates the genetic traits of files by using a proprietary algorithm. These traits define the potential of the software to carry out malicious or harmless actions when executed on a computer. GHE is capable of determining whether a file is innocuous, worm, spyware, Trojan, or virus 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. Sensitivity levels are applied to different environments depending on the probability of malware prevalence. For example, at network SMTP gateways, the likelihood of executable files being malware is very high. Therefore, implementations in commercial products are set at high sensitivity for network layer email scanning products. However for storage or application layers where the vast majority of executable code is from legitimate applications, GHE can be implemented with medium sensitivity. This setting provides the ability to maximize detection rates for unknown malware while having a negligible false positive rate. The results obtained with GHE have been excellent. Since its release, roughly one third of the new variants received at Panda Security from real users' machines have been submitted automatically by GHE.

## *Fourth-Generation Security Solutions: Collective Intelligence*

Malware is being distributed at a rate 10 times greater than two years ago. Security solutions must now detect 10 times more malware to provide adequate protection to users. While a full-fledged HIPS solution raises the bar substantially by detecting and blocking most of these with proactive technologies, it is still possible for unknown malware to slip through its defenses. Consider the fact that while 80-90% of proactive effectiveness could be considered an excellent score – in absolute terms it may lead to hundreds or thousands of samples being missed over time, since even a small fraction of a big number will equate to a huge amount of malware.

Panda Security's Collective Intelligence solution was released at the end of 2006 with the objective of being able to reliably detect 10 times more malware than they were then detecting, with 10 times less effort. Collective Intelligence functions as an online and real-time Security-as-a-Service (SaaS) platform. Now, with over two years of research and development and millions of dollars in investment, the Collective Intelligence approach provides tremendous value by:

1. Leveraging community knowledge to proactively protect all users
2. Automating and enhancing malware collection, classification, and remediation
3. Gaining knowledge on techniques to improve existing technologies
4. Deploying new generation of security services "from the cloud"

### Benefiting from Community Knowledge

Traditional security solutions are typically 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. As a result, 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 in different PCs. Most importantly, other PCs do not automatically benefit from proactive malware detections on different PCs. They have to wait for the antivirus lab to receive that specific sample, for a signature to be created, QA'd, and deployed to ultimately protect other users.

As a result, traditional approaches are just simply too slow to combat today's rapidly moving malware. One of the main benefits of Panda Security's Collective Intelligence approach – in addition to the effectiveness provided by the automation of the malware remediation lifecycle – is the automatic and real-time benefit it provides to the users of the Collective Intelligence community. As soon as a malicious process is detected in a users' PC by the Collective Intelligence servers (whether by system heuristics, emulation, sandboxing, or behavioral analysis, etc.) the rest of the users worldwide automatically benefit from that specific detection. This results in close to real-time detection – not only of initial malware outbreaks – but also of targeted attacks whose objective is infecting a small number of users to stay below the radar.

### The Automated Malware Protection Process

One of the biggest barriers to raising the bar of reliable malware detection ratios is the process of creating a signature against a single sample takes too long. Each malware sample needs to be sent to the lab by an affected user or fellow researcher, reversed engineered by a lab technician to create a detection signature and disinfection routine. The results then need to be QA'd, uploaded to production servers, replicated worldwide, and finally downloaded and applied by customers. This entire process

is mostly manual and can take up anywhere from minutes to hours or days or even weeks, depending on the workload of the lab engineers and other factors, such as sample priority, prevalence, damage potential, media coverage, etc.

The process can be delayed much longer when intelligence or functionality upgrades to the anti-malware or behavioral engines are involved. Typically anti-malware vendors upgrade their solutions once or twice a year, as each upgrade has a costly testing and deployment process for customers. With Panda Security's Collective Intelligence infrastructure, this entire process of malware collection, classification, and remediation can be automated and routinely performed online on a real-time basis for the vast majority of samples.

The following section will walk through the malware protection process from the point of view of a computer that has just been exploited and infected by malicious code.

***1. Automated Malware Collection***

The Panda Collective Intelligence agent gathers information on processes and memory objects and performs queries against the CI central servers which perform a variety of checks against them. If certain conditions are met, the suspicious file (or part thereof) is automatically uploaded with the users consent to the CI servers where it is further processed.

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. Panda Security has built a vast database of malware samples which are automatically collected, which in turn provide the CI web-service with a real-time feed of new malware classification entries.

***2. Automated Malware Classification***

Server-based processing is not limited by the CPU and memory constraints of the PCs. Therefore scanning routines at the CI servers undergo much more in-depth processing by more sensitive technologies (signature and sensitive heuristics scanning, emulation, sandboxing, virtualization, white-listing, 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 typical scenarios at PCs, desktops, or server machines. Therefore, many of the more resource-intensive proactive techniques which Panda Security is using and which provide much higher detection rates can now be used for the benefit of the entire user-base without even touching customers' valuable 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 Panda Security and therefore samples that cannot be classified automatically are ultimately looked at by an analyst at the lab.

***3. Automated Malware Remediation***

The remediation module of the CI 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 entire community of CI users to proactively detect and disinfect new or even targeted attacks with very low numbers of infected hosts.

Traditional anti-malware and HIPS solutions have already started to benefit from the CI approach. During the initial three months of its operation, the Panda Security remediation module has created protection for hundreds of thousands of malware samples which have been gradually deployed to

the company's existing products.

One of the main benefits of the Collective Intelligence approach is that signatures do not need to be downloaded to each client as they operate "from the cloud". This, however, does not mean that the client machine will not need to maintain updated signatures. A potential threat to such an approach is the availability of the Collective Intelligence servers. However, Panda Security's Collective Intelligence technology is designed as an additional layer of protection, integrated with any other of the user's current solutions. Therefore, under non-availability of the CI platform for whatever reason, security protection would fall back to the user's regular HIPS solution, providing above average protection.

## Gaining Knowledge on Malware Techniques

One of the main benefits provided by the community feature of Collective Intelligence is that it provides continual insight for Panda Security's engineers into new malware techniques and distribution points. Questions such as where a specific piece of malware was first found and how it spread enable Panda Security to model additional intelligence into specific malware families and even sometimes to identify creators of specific malware variants.

The approach of applying data warehousing and data mining techniques to malware detection 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 applied for tracking infection origins, which in turn has interesting applications and benefits for law enforcement efforts.

## *Deploying Security Services "From the Cloud"*

Panda Security has developed and deployed several 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 consumers and stand-alone PCs, Panda NanoScan (www.nanoscan.com) scans a PC for malware actively running, and TotalScan (www.pandasecurity.com/totalscan) automatically performs a full system scan of the entire PC including hard drive, memory, email databases, etc.

The requirements for performing an in-depth malware audit are more demanding for corporate environments. To address the needs of enterprises of all sizes, Panda Security has 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, and unveil machines which have already been subject to targeted attacks.

# Conclusion

The latest advances by the cyber-crime communities are taking advantage of the inherent weaknesses in the security industry. Labs are being swamped by a plethora of malware being created continually. But, because it remains invisible, most users do not perceive the need for additional protection. Targeted attacks that only infect a few users are more effective than epidemic attacks that infect millions of users. Unfortunately, the majority of users tend to trust a single solution or single layer of protection as their main line of defense against malware. The need for additional protection is revealed by studies that show that a large portion of users with current and updated security solutions are in fact infected.

To tackle these cyber-crime advances, new layers of protection are needed that take advantage of automating the entire malware protection cycle – from sample collection, analysis, classification, to remediation. But automation by itself is not enough. Enterprises need visibility into what is happening on all PCs in order to detect targeted attacks more quickly and efficiently and gain a competitive edge on malware creators.

Collective Intelligence – the approach developed by Panda Security – gives all the benefits of an added layer of defense. It provides effective response and protection to malware threats, is able to detect targeted attacks, and gains intelligence thanks to the correlation of all the detections by the community of users. In short, it takes a planet to protect one PC!

For more information on Panda Security's solutions and the Collective Intelligence approach to fighting malware, please visit www.pandasecurity.com.

# Glossary

**Administrator**: a person or program responsible for managing and monitoring an IT system or network, assigning permissions, etc.

**Adware**: programs that display advertising using any means: pop-ups, banners, changes to the browser home page or search page, etc. It is sometimes installed with the user's consent and knowledge, but on other occasions it is not. It operates in the same way regardless of whether the user has consented or not.

**Algorithm**: a process or set of rules

**Antivirus / antivirus programs**: these are programs that scan the memory, disk drives and other parts of a computer for viruses.

**API (Application Program Interface)**: this is a function used by programs to interact with operating systems and other programs.

**Backdoor**: this is a point (hardware or software) through which it is possible to control the affected system totally or partially without the user realizing.

**Boot / Master Boot Record (MBR)**: Also known as the Boot sector, this is the area or sector of a disk that contains information about the disk itself and its properties for starting up the computer.

**Boot virus**: a virus that specifically affects the boot sector of both hard disks and floppy disks.

**Bot**: a term derived from 'robot'. This is a Trojan or worm designed to install itself on computers and take action automatically when it receives commands from a remote attacker.

**Bot Herder (or Bot Master)**: name for the person that controls the botnets.

**Botnet**: this is a group of bots which are interconnected and which thereby, illegally, connect a group of computers.

**Browser**: a browser is the program that lets users view Internet pages. . The most common browsers are: Internet Explorer, Netscape Navigator, Opera, etc.

**Cleaning**: the action that an antivirus takes when it detects a virus and eliminates it.

**Client**: IT system (computer) that requests certain services and resources from another computer (server), to which it is connected across a network.

**Code**: content of virus files -virus code- written in a certain programming language. It can also refer to systems for representing or encrypting information. In its strictest sense, it can be defined as a set of rules or a combination of symbols that have a given value within an established system.

**Communications port**: point through which a computer is accessed and information is exchanged (inbound/outbound) between the computer and external sources (via TCP/IP).

**Converged HIPS**: these are HIPS integrating different technologies against different types of threats. The technologies included can be both reactive and proactive.

**Denial of Service (DoS)**: This is a type of attack, sometimes caused by viruses, that prevents users from accessing certain services (in the operating system, web servers etc.).

**Dialer**: this is a program that is often used to maliciously redirect Internet connections while browsing the Internet. When used in this way, they disconnect the legitimate telephone connection used to hook up to the Internet and re-connect it via a premium-rate number. This activity results in an extremely expensive phone bill.

**Disinfection**: the action that an antivirus takes when it detects a virus and eliminates it.

**DNS (Domain Name System)**: system that translates internal numeric Internet addresses into comprehensible names. All operations and communications carried out on the Internet are based on providing addresses. However, internally, the Internet actually uses numeric addresses to enable communications. For example: 127.163.24.3. DNS is the system that translates numeric addresses into alphanumeric addresses and vice-versa.

**Exploit**: technique or program that exploits a security flaw -a vulnerability- in a certain communication protocol, operating system or IT tool.

**Firewall**: this is a process or program that monitors and scans traffic entering or exiting a computer from a network, in all protocols and on every port, in order to prevent threats like network viruses, hackers and others from entering.

**Gateway**: a computer that allows communication between different types of platforms, networks, computers or programs. To do this it translates the different communications protocols they use. It is what is also known as the access point.

**Genetic Heuristic Engine**: New generation of heuristic detection developed by Panda Software to determine malware presence through so-called 'genetic scanning'. Compared to traditional heuristics, genetic heuristics looks at many more aspects to determine whether a file contains malware or not with maximum precision. Genetic scanning detection does not simply compare patterns, but uses a great variety of parameters.

**Hacker**: someone who accesses a computer illegally or without authorization.

**Heuristic scan**: the term 'heuristics' refers to problem solving by trial and error. Used in the computer world, it refers to a technique used for detecting unknown viruses.

**HIPS (Host Intrusion Prevention System)**: See IPS.

**Honeypots**: Systems connected to the Internet used as a bait for hackers but which do not contain useful information. They are used for detecting vulnerabilities or failures in protection systems.

**IDS (Intrusion Detection System)**: When an IT system connects to the Internet, it actually connects to a network where there are hundreds of millions of users at the same time. Any of them could (and many of them do) enter the computer illegally. To prevent this, there are many solutions on the market for detecting intrusion attempts to systems. These are called IDS. However, this type of protection is not sufficient, as it requires intervention from network administrators to distinguish between false alarms and real threats and take the necessary corrective action manually.

**In The Wild**: this is an official list drawn up every month of the viruses reported causing incidents.

**IPS**: Compared to IDS, the IPS (Intrusion Prevention System) not only detects unknown threats more accurately, but also can take preventive actions to stop threats from spreading and damaging systems. Depending on their location, IPS are divided into:

**HIPS (Host Intrusion Prevention System)**: This system type watches many sources of suspicious activity such as calls from processes to the operating system, communications with other computers, or memory usage, HIPS are installed to detect malicious use of computers from within the network and from users who might have managed to by-pass traditional protection systems such as corporate antiviruses and firewalls.

**NIPS (Network Intrusion Prevention System):** These systems offer network perimeter protection against intrusions acting as an advanced firewall, that is, inspecting network packets or attempts to exploit system vulnerabilities through remote calls. Most NIPS are implemented in organizations through high-performance appliances to avoid loss in the quality of the service.

**PIPS (Personal Intrusion Prevention System):** These are systems that prevent intrusions in home computers and home offices, as well as workstations in companies.

**IT resources**: the IT components that belong to an organization.

**Jokes**: these are not viruses, but tricks that aim to make users believe they have been infected by a virus.

**Kernel**: this is the central module of an operating system.

**Keylogger**: a program that collects and saves a list of all keystrokes made by a user in order to obtain confidential information.

**LAN (Local Area Network)**: a network of interconnected computers in a reasonably small geographical area (generally in the same city or town or even building).

**Macro**: a macro is a series of instructions defined so that a program, say Word, Excel, PowerPoint, or Access, carries out certain operations. As macros are programs, they can be affected by viruses. A virus that affects macros in files is called macro virus.

**Macro virus**: this is a virus that affects the macros contained in Word documents, Excel spreadsheets, PowerPoint presentations, etc.

**Malware**: this term is used to refer to all programs that contain malicious code (MALicious softWARE), whether it is a virus, Trojan or worm.

**Network**: a group of computers or devices which are interconnected with cables, telephone lines, or electromagnetic waves (microwaves, satellite, etc.) in order that they can communicate and share resources. The Internet is an enormous network to which other sub-networks with millions of computers are connected.

**Network perimeter**: the components in a network that can be connected to other networks or users outside the network.

**NIPS (Network Intrusion Prevention System)**: See IPS.

**Password**: this is a sequence of characters used to restrict access to a certain file, program or other area, so that only those who know the password can enter. A typical example is the credit card password.

**Payload**: the effects of a virus.

**Pharming**: an online fraud technique that alters the DNS so that the victim, when they enter a certain Internet address, is taken to a spoof site designed to steal confidential data. Pharming can involve modifying the DNS server itself, thereby affecting many users at the same time, or locally modifying the Windows HOSTS file on an individual computer.

**Phishing**: online fraud technique that involves directing users to a false web page to steal personal data. In general, spoof emails are used that appear to come from reliable sources.

**PIPS (Personal Intrusion Prevention System)**: See IPS.

**Polymorphic / Polymorphism**: a technique used by viruses to encrypt their signature in a different way every time and even the instructions for carrying out the encryption.

**Proactive detection technologies**: technologies that can detect a malware sample by themselves, not needing to know it beforehand.

**Proactive technology**: in IT security these are technologies that are able to act by themselves, without the need to use virus signatures or other types of updates.

**Proxy**: a proxy server acts as an intermediary between an internal network (an intranet, for example) and an external connection to the Internet. This allows a connection for receiving files from web servers to be shared.

**Reactive detection technologies**: technologies that can detect only malware that has been previously identified. A clear example is detection using signature files. An antivirus that uses this technology will only detect a new threat if it includes an updated signature file that contains the vaccine against it. Otherwise, it won't detect it.

**Rootkit**: a rootkit is a set of tools used frequently by intruders or crackers that illicitly access an IT system. These tools enable them to hide the processes and files that allow the intruder to access the system, often with malicious intent. There are rootkits for many kinds of operating systems, such as Linux, Solaris or Microsoft Windows.

**Script / Script virus**: the term script refers to files or sections of code written in programming languages like Visual Basic Script (VBScript), JavaScript, etc.

**Security patch**: set of additional files applied to a software program or application to resolve certain problems, vulnerabilities or flaws.

**Server**: IT system (computer) that offers certain services and resources (communication, applications, files, etc.) to other computers (known as clients), which are connected to it in a network.

**Service**: the suite of features offered by one computer or system to others that are connected to it.

**Signature / Identifier**: this is like the virus passport number. A sequence of characters (numbers, letters, etc.) that identify the virus.

**SMTP (Simple Mail Transfer Protocol)**: this is a protocol used on the Internet exclusively for transfering email messages. This protocol is used to send email through the Internet or to connect non-compatible email servers. To access the email stored in a user's mailbox, the following protocol is used: POP (Post Office Protocol). This is due to the fact that SMTP doesn't have authentication systems.

**Sniffer**: a program that looks for numeric or character strings in packets that pass through network nodes in order to get information.

**Sniffing**: capture of packets through a sniffer.

**Software**: files, programs, applications and operating systems that enable users to operate computers or other IT systems. These are the elements that make the hardware work.

**Spam**: massive sending and receiving of unsolicited email.

**Spear Phishing**: this is a targeted phishing attack. In this case, instead of the usual phishing message sent to millions of users, it is sent to a small group of users, such as members of a company or institution.

**Split Tunneling**: IPSec VPNs that enable splitting communication traffic according to its destination: the VPN server (corporate traffic) or Internet servers (normal traffic).

**Spyware**: programs designed to be installed silently on a system and steal data related to users' Internet movements.

**Stealth**: a technique used by viruses to infect computers unnoticed by users or antivirus applications (temporarily).

**Targeted attack**: IT attacks aimed at a small group of users. There are two types: indiscriminate, which are aimed at a small, but random, group of users, or specific, which target a specific user or company.

**Trojan/ Trojan horse**: strictly speaking, a Trojan is not a virus, although it is often thought of as such. This is a program that enters computers (in a number of ways), installs itself and carries out actions that enable them to take control of the affected computer. Its name is inspired by the Trojan Horse story from Greek mythology.

**Update / Updates**: antiviruses are constantly becoming more powerful and adapting to the new technologies used by viruses. If they are not to become obsolete, they must be able to detect the new viruses that are constantly appearing. To do this, they have what is called a Virus Signature File. This file includes all the characteristics that identify viruses so that they can be detected and the appropriate action can be taken. The process of incorporating the latest version of this file and other files in the antivirus software is known as an update.

**Variant**: a variant is a modified version of an original virus, which may vary from the original in terms of means of infection and the effects that it has.

**VPN (Virtual Private Network)**: Permanent virtual network for guaranteeing encryption of data transferred between a local network and the Internet. Today, many companies provide employees with mobile communication systems, like laptops for example. In this way they can work out of the office and

still be able to transmit results or get new data from the company systems. This communication usually takes place through the Internet, which is a public system and could allow a hacker to intercepted it and spy on data transferred. To avoid this, data encryption and decryption systems are installed on the company and the employee's laptop, or on two segments of the company's network. In this way, information travels through a public channel, but cannot de decoded by unauthorized users. It is as if the network was a private network mounted on a public network. Hence, the name Virtual Private Network.

**WiFi (Wireless Fidelity)**: High-frequency wireless local area network.

**Windows registry**: This is a file that stores all configuration and installation information of programs installed, including information about the Windows operating system.

**Worm**: a program that is similar to a virus, but differs in that all it does is make copies of itself or part of itself.

**WLAN (Wireless LAN)**: Local area network with wireless connections.

**Zero day**: Exploit designed for a vulnerability for which there is no patch available yet. Generally zero-day exploits are released on the same day the targeted vulnerability is discovered.