SE LABS ®

EDR
DETECTION
Wizard Spider

# Advanced Security Test Report

## **AhnLab** EPP/EDR

ONLINE REPORT

NOVEMBER 2025

SE LABS ® tested **AhnLab EPP/EDR** against a range of hacking attacks designed to compromise systems and penetrate target networks in the same way as criminals and other attackers breach systems and networks.

Full chains of attack were used, meaning that testers behaved as real attackers, probing targets using a variety of tools, techniques and vectors before attempting to gain lower-level and more powerful access. Finally, the testers/attackers attempted to complete their missions, which might include stealing information, damaging systems and connecting to other systems on the network.

# Contents




Document version 1.0 Written 5th November 2025




SE LABS




**Website** selabs.uk
**Email** info@SELabs.uk
**LinkedIn** www.linkedin.com/company/se-labs/
**Blog** selabs.uk/blog
**Post** SE Labs Ltd,
55A High Street,
Wimbledon,
SW19 5BA, UK



SE Labs is ISO/IEC 27001 : 2022 certified and BS EN ISO 9001 : 2015 certified for The Provision of IT Security Product Testing.

CEO
**Simon Edwards**

If you spot a detail in this report that you don't understand, or would like to discuss, please contact us. SE Labs uses current threat intelligence to make our tests as realistic as possible. To learn more about how we test, how we define 'threat intelligence' and how we use it to improve our tests please visit our website and follow us on LinkedIn.

# Endpoint Detection and Response is more than anti-virus

## Gain insights into cyber security testing through transparent threat intelligence.

**An Endpoint Detection and Response** (EDR) product goes beyond traditional antivirus software, which is why it requires more sophisticated testing. This involves testers mimicking real attackers and following every step of an attack.

While shortcuts might seem tempting, fully executing each phase of an attack is crucial to truly evaluate the effectiveness of EDR products.

Moreover, each step must reflect real-world scenarios; you can't just guess what cybercriminals might do and hope it's accurate. That's why SE Labs tracks the actual behaviour of cyber criminals and designs tests based on how attackers attempt to compromise their targets.

The cyber security industry refers to this sequence of steps as the 'attack chain.' The MITRE organization has documented these stages in its ATT&CK framework.

While this framework doesn't provide an exact blueprint for real-world attacks, it offers a structured guide that testers, security vendors, and customers (like you!) can use to conduct tests and interpret the results.

SE Labs' Enterprise Advanced Security tests are based on real attacker behaviour, and we present our findings using a MITRE ATT&CK-style format.

You can see how the ATT&CK framework outlines each step of an attack and how we apply it to our testing in section **4. Threat Intelligence**, starting on page 12. This approach offers two key benefits: confidence that our tests are both realistic and relevant, and familiarity with the way cyber attacks are illustrated.

# Executive Summary

SE Labs tested **AhnLab EPP/EDR** against targeted attacks based on Wizard Spider, a threat group known to be associated with prolific malware attacks.

**We examined its abilities to:**

- **Detect highly targeted attacks**
- **Provide remediation to damage and other risks posed by the threats**
- **Handle legitimate applications and other objects**

Legitimate files were used alongside the threats to measure any false positive detections or other sub-optimal interactions.

**AhnLab EPP/EDR** scored a perfect 100% Detection Accuracy Rating. It detected the delivery and execution of the Wizard Spider attacks.

The product detected all the subsequent malicious activities in the attack chain, tracking all of the hostile activities that occurred as the attacks progressed.

**AhnLab EPP/EDR** only misclassified a few legitimate objects, earning a Legitimacy Accuracy Rating of 90%.

Given its Total Accuracy Rating of 95%, the product can be described as very accurate and achieved the AAA rating for Advanced Security EDR Detection.

## Executive Summary

| **AhnLab EPP/EDR** | | | |
|---|---|---|---|
| | Detection Accuracy Rating (%) | Legitimate Accuracy Rating (%) | Total Accuracy Rating (%) |
| AhnLab EPP/EDR | 100% | 90% | 95% |

- **Products highlighted in green were the most accurate, scoring 90 per cent or more for Total Accuracy. Those in orange scored less than 90 but 71 or more. Products shown in red scored less than 71 per cent.**

For exact percentages, see **2. Total Accuracy Ratings** on page 9.

## Advanced Security Test Award

The following product wins the SE Labs award:


**AhnLab**
EPP/EDR

# 1. How We Tested

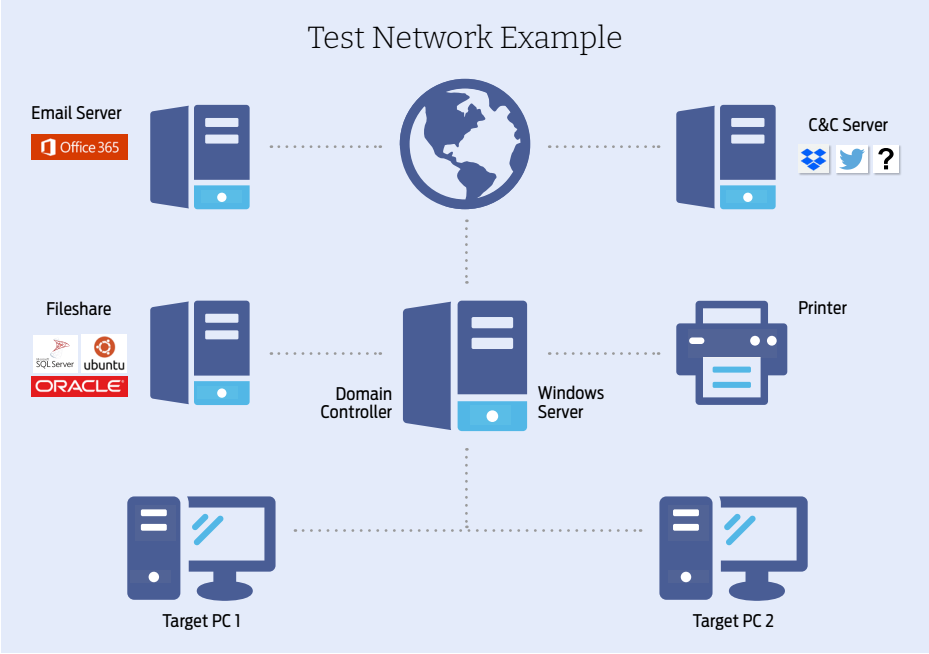
**Testers can't assume** that products will work a certain way, so running a realistic advanced security test means setting up real networks and hacking them in the same way that real adversaries behave.

In the diagram on the right you will see an example network that contains workstations, some basic infrastructure such as file servers and a domain controller, as well as cloud-based email and a malicious command and control (C&C) server, which may be a conventional computer or a service such as Dropbox, Twitter, Slack or something more imaginative.

As you will see in the **Threat Responses** section on page 7, attackers often jump from one compromised system to another in so-called 'lateral movement'. To allow products to detect this type of behaviour the network needs to be built realistically, with systems available, vulnerable and worth compromising.

It is possible to compromise devices such as enterprise printers and other so-called 'IoT' (internet of things) machines, which is why we've included a representative printer in the diagram.

The techniques that we choose for each test case are largely dictated by the real-world behaviour of online criminals. We observe their tactics and replicate what they do in this test. To see more details about how the specific attackers behaved, and how we copied them, see **Attack Details** on page 8 and, for a really detailed drill down on the details, **4. Threat Intelligence** on page 12 and **Appendix D: Attack Details** on page 18.




● This example of a test network shows one possible topology and ways in which enterprises and criminals deploy resources

# Threat Responses

**Full Attack Chain: Testing Every Layer of Detection and Protection**

Attackers start from a certain point and don't stop until they have either achieved their goal or have reached the end of their resources (which could be a deadline or the limit of their abilities). This means that, in a test, the tester needs to begin the attack from a realistic first position, such as sending a phishing email or setting up an infected website, and moving through many of the likely steps leading to actually stealing data or causing some other form of damage to the network.

If the test starts too far into the attack chain, such as executing malware on an endpoint, then many products will be denied opportunities to use the full extent of their protection and detection abilities. If the test concludes before any 'useful' damage or theft has been achieved, then similarly the product may be denied a chance to demonstrate its abilities in behavioural detection and so on.

## Attack Stages

The illustration (below) shows typical stages of an attack. In a test, each of these should be attempted to determine the security solution's effectiveness. This test's results record detection and protection for each of these stages.

We measure how a product responds to the first stages of the attack with a detection and/or protection rating. Sometimes products allow threats to run yet still detect them. Other times they might allow the threat to run briefly before neutralising it. Ideally, they detect and block the threat before it has a chance to run. Products may delete threats or automatically contain them in a 'quarantine' or other safe holding mechanism for later analysis.

Should the initial attack phase succeed, we then measure post-exploitation stages, which are represented by steps two through to seven below. We broadly categorise these stages as: Access (step 2); Action (step 3); Escalation (step 4); and Post-Escalation (steps 5-6).

**In figure 1.** you can see a typical attack running from start to end, through various 'hacking' activities. This can be classified as a fully successful breach.

**In figure 2.** a product or service has interfered with the attack, allowing it to succeed only as far as stage 3, after which it was detected and neutralised. The attacker was unable to progress through stages 4 onwards.

It is possible for an attack to run in a different order with, for example, the attacker attempting to connect to other systems without needing to escalate privileges. However, it is common for password theft (see step 5) to occur before using stolen credentials to move further through the network.


**Figure 1.** A typical attack starts with an initial contact and progresses through various stages, including reconnaissance, stealing data and causing damage.


**Figure 2.** This attack was initially successful but only able to progress as far as the reconnaissance phase.

# Attack Details

**When testing services** against targeted attacks it is important to ensure that the attacks used are relevant. Anyone can run an attack randomly against someone else. It is the security vendor's challenge to identify common attack types and to protect against them. As testers, we need to generate threats that in some way relate to the real world.

All of the attacks used in this test are valid ways to compromise an organisation. Without any security in place, all would succeed in attacking the target. Outcomes would include systems infected with ransomware, remote access to networks and data theft.

But we didn't just sit down and brainstorm how we would attack different companies. Instead we used current threat intelligence to look at what the bad guys have been doing over the last few years and copied them quite closely. This way we can test the services' abilities to handle similar threats to those faced by global governments, financial institutions and national infrastructure.

| Attacker/ APT Group | Method | Target | Details |
|---|---|---|---|
| **Wizard Spider** | Phishing Attachment |  | Credential harvesting, cryptomining and implementation of ransomware. |

| KEY | | | | | |
|---|---|---|---|---|---|
|  | Education |  | Financial Industries |  | Gambling |
|  | Government Espionage |  | Manufacturing |  | Natural Resources |
|  | Private-sector Energy |  | Research Institutes |  | Travel Industries |

The graphic on this page shows a summary of the attack group that inspired the targeted attacks used in this test. If a service was able to detect and protect against these then there's a good chance they are on track to blocking similar attacks in the real world. If they fail, then you might take their bold marketing claims about defeating hackers with a pinch of salt.

For more details about each APT group please see **4. Threat Intelligence** on page 12.

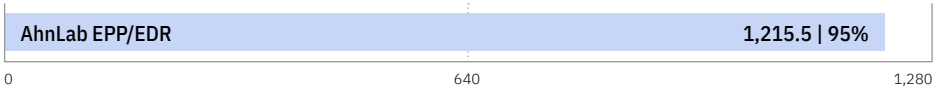
# 2. Total Accuracy Rating

**This test examines** the total insight a product has, or can provide, into a specific set of attacking actions. We've divided the attack chain into chunks of one or more related actions. To provide sufficient insight, a product must detect at least one action in each chunk.

If you look at the results tables in **Response Details** on page 11 you'll see that Delivery and Execution are grouped together into one chunk, while Action sits alone. Escalation and Post-Escalation (PE) Action are grouped, while Lateral Movement and Lateral Action are also grouped.

This means that if the product detects either the threat being delivered or executed, it has coverage for that part of the attack. If it detects the action as well as the escalation of privileges and an action involved in lateral movement then it has what we consider to be complete insight, even if it doesn't detect some parts of some chunks (i.e. Lateral Movement, in this example).

## Total Accuracy Rating




- **Total Accuracy Ratings combine protection and false positives.**

# 3. Response Details

**In this test** security products are exposed to attacks, which comprise multiple stages. The perfect product will detect all relevant elements of an attack. The term 'relevant' is important, because sometimes detecting one part of an attack means it's not necessary to detect another.

For example, in the table below certain stages of the attack chain have been grouped together. As mentioned in **2. Total Accuracy Ratings**, these groups are as follows:

### Delivery/Execution (+10)

If the product detects either the delivery or execution of the initial attack stage then a detection for this stage is recorded.

### Action (+10)

When the attack performs one or more actions, while remotely controlling the target, the product should detect at least one of those actions.

### Privilege Escalation/Action (+10)

As the attack progresses there will likely be an attempt to escalate system privileges and to perform more powerful and insidious actions. If the product can detect either the escalation process itself, or any resulting actions, then a detection is recorded.

### Lateral Movement/Action (+10)

The attacker may attempt to use the target as a launching system to other vulnerable systems.

If this attempt is discovered, or any subsequent action, a detection is reported.

The Detection Rating is calculated by adding points for each group in a threat chain that is detected. When at least one detection occurs in a single group, a 'group detection' is recorded and 10 points are awarded. Each test round contains one threat chain, which itself contains four groups (as shown below), meaning that complete visibility of each attack adds 40 points to the total value.

A product that detects the delivery of a threat, but nothing subsequently to that, wins only 10 points, while a product that detects delivery and action, but not privilege escalation or lateral behaviours, is rated at 20 for that test round.

## Understanding Detection Groups

| | | First Group | | Second Group | Third Group | | Fourth Group | |
|---|---|---|---|---|---|---|---|---|
| Incident No: | Detection | Delivery | Execution | Action | Escalation | PE Action | Lateral Movement | Lateral Action |
| 1 | ✓ | ✓ | ✓ | — | ✓ | ✓ | ✓ | ✓ |
| 2 | ✓ | — | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| 3 | ✓ | — | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| 4 | ✓ | ✓ | ✓ | — | ✓ | ✓ | ✓ | ✓ |

| Attacker/ Apt Group | Number of Incidents | Attacks Detected | Delivery/ Execution | Action | Privilege Escalation/Action | Lateral Movement Action |
|---|---|---|---|---|---|---|
| Dragonfly & Dragonfly 2 | 4 | 4 | 4 | 2 | 4 | 4 |

Elements of the attack chain are put into groups. For example, the Delivery and Execution stages of an attack are in the same group. Similarly, we group the Post Escalation stage with the Post Escalation Action (PE Action) stage. When we count detections we look to see at least one detection (tick) in each group. One or two detections in a group is a success.

In this example we have four test cases, which we call 'incidents'. In Incident No. 1 there was a detection recorded for the delivery of the threat and when it was executed. These two results count as one detection. In Incident No. 2 the threat delivery was not detected, but its execution was. This also counts as one detection.

When no detection is registered in any part of a group the result will be a 'miss'. In Incident 1. there was no detection when the attacker performed the 'Action' stage of the attack. This is a miss for the product. In fact, this product only detected two of the four Action stages, which is why the Response Details table shows '2' in the Action column.

## Wizard Spider

| Incident No: | Detection | Delivery | Execution | Action | Escalation | PE Action | Lateral Movement | Lateral Action |
|---|---|---|---|---|---|---|---|---|
| 1 | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | — |
| 2 | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| 3 | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| 4 | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |

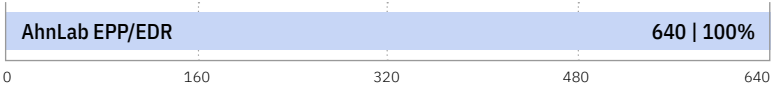
## Response Details

| Attacker/ APT Group | Number of Incidents | Attacks Detected | Delivery/ Execution | Action | Privilege Escalation/ Action | Lateral Movement/ Action |
|---|---|---|---|---|---|---|
| Wizard Spider | 4 | 4 | 4 | 4 | 4 | 4 |
| **TOTAL** | **4** | **4** | **4** | **4** | **4** | **4** |

## Detection Accuracy Rating Details

| Attacker/APT Group | Number of Incidents | Attacks Detected | Group Detections | Detection Rating |
|---|---|---|---|---|
| Wizard Spider | 4 | 4 | 16 | 160 |
| **TOTAL** | **4** | **4** | **16** | **160** |

## Detection Accuracy Rating




- Detection Ratings are weighted to show that how products detect threats can be subtler than just 'win' or 'lose'.

## Group Detections

We record detections in groups, as described above in Understanding Detection Groups. To get an overview of how a product handled the entire set of threats we then combine these detections into 'Group Detections'.

In a test with four incidents and four detection groups (Delivery/Execution; Action; Escalation/PE Action; and Lateral Movement/Lateral Action) the maximum score would be 16. This is because for each of the four threats a product that detects everything would score 4.

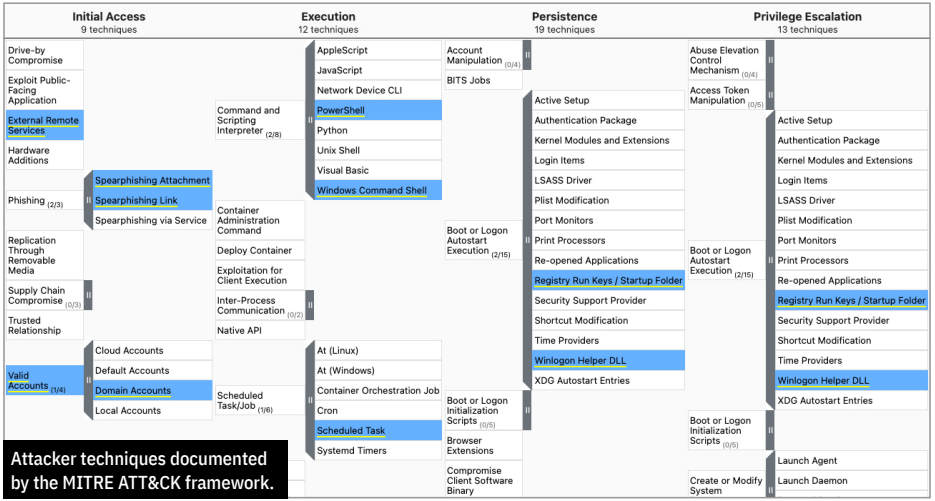
Our overall Detection Rating is based on the number of Detection Groups achieved.

# 4. Threat Intelligence

## Wizard Spider

Known to have operated since at least 2016, Wizard Spider is considered to be a threat group based in and around St. Petersburg, Russia. It is most notable for developing the TrickBot banking malware. Wizard Spider has infected over a million systems worldwide predominantly by using this malware.

**Reference:**
https://attack.mitre.org/groups/G0102/




### Example Wizard Spider Attack

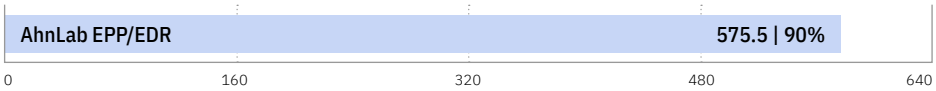
| Delivery | Execution | Action | Privilege Escalation | Post-Escalation | Lateral Movement | Lateral Action |
|---|---|---|---|---|---|---|
| Spear Phishing Attachment | Windows Command Shell | File and Directory Discovery | Bypass User Account Control | Remote System Discovery | Service Execution | Archive Collected Data |
| | Malicious File | Process Discovery | Valid Accounts | Security Software Discovery | Domain Accounts | Data Staged |
| | Obfuscated Files or Information | System Information Discovery | | LLMNR/NBT-NS Poisoning and SMB Relay | | Data from Local System |
| | Powershell | System Network Configuration Discovery | | | | Exfiltration Over C2 Channel |
| | | System Owner/User Discovery | | | | |

# 5. Legitimate Accuracy Rating

**These ratings indicate** how accurately the product classifies legitimate applications and URLs, while also taking into account the interactions that the product has with the user. Ideally a product will either not classify a legitimate object or will classify it as safe. In neither case should it bother the user.

We also take into account the prevalence (popularity) of the applications and websites used in this part of the test, applying stricter penalties for when products misclassify very popular software and sites.

## Legitimate Accuracy Rating




- **Legitimate Accuracy Ratings can indicate how well a vendor has tuned its detection engine.**

# 6. Conclusion

This test exposed **AhnLab EPP/EDR** to a set of Wizard Spider exploits, file-less attacks and malware attachments. These diverse attacks were delivered through spear phishing links in the same way that the Wizard Spider threat has been witnessed in real-world attacks over the previous few years. They are representative of a real and present threat to government and business networks the world over.

The threats used in this test are similar or identical to those used by the threat group described in **Attack Details** on page 8 and **Threat Intelligence** on page 12. This Russia-based group has operated since at least 2016, although its tactics have evolved over time. SE Labs has monitored and adapted its Wizard Spider attacks to match the current reality.

It is important to note that while the test used the same type of attacks, new files were used. This exercised the tested product's abilities to detect and protect against certain approaches to attacking systems rather than simply detecting malicious files that have become well-known over the previous few years. The results are an indicator of future performance rather than just a compliance check that the product can detect old attacks.

**AhnLab EPP/EDR** detected all of the threats on a basic level, in that for each attack it detected at least some elements of the attack chain. It detected all of the elements of the first Detection Group, issuing alerts for all of the spear phishing messages as well as the variety of executables to which these messages were linked.

Impressively, the product also detected all of the actions in the second Detection Group. It issued alerts when files and directories, or information about the system or the system owner were being exposed. It also detected attacker control over local data and exfiltration over the C2 channel in every case.

**AhnLab EPP/EDR's** performance was similarly excellent when tested against the elements of the third Detection Group. It detected every attempt to escalate privileges and post-escalation actions, catching attempts to gain privileges over domain accounts as well as attempts to bypass user account control.

The product was also awarded full scores for detecting the attacks as they moved from the target system to the other devices in the network. It missed only one instance of lateral action.

The product scored a 100% Detection Accuracy Rating. This is offset by its 90% Legitimate Accuracy Rating, given for correctly classifying almost all of the legitimate objects in the test. **AhnLab EPP/EDR** achieved a Total Accuracy Rating of 95% and was awarded with the AAA rating for Advanced Security EDR Detection.

# Appendices

## Appendix A: Legitimate Interaction Ratings

**It's crucial that** security products not only detect threats but also correctly handle legitimate objects, such as files and URLs. Incorrectly labelling legitimate objects as being 'malware' or 'harmful' is a false positive (FP) result.

In reality, genuine FPs are quite rare in good testing, with good products. In our experience it is unusual for a legitimate application to be classified as 'malware'. More often it will be classified as 'unknown', 'suspicious' or 'unwanted' (or other terms that mean much the same thing).

### Interaction Ratings

We use a subtle system to rate a product's approach to legitimate objects. This takes into account how it classifies them and how it presents that information.

Sometimes a product will pass the buck and demand that a user or administrator decide if something is safe or not. In such cases, the product may make a recommendation to allow or remove the object. In other cases the product will make no recommendation, which is possibly even less useful.

If a product reports that an application is safe, or doesn't recommend any action (such as to remove it), it has achieved an optimum result. Anything else is a Non-Optimal Classification/Action (NOCA).

A product may be configured with a policy to restrict certain objects according to the business' objectives. A recommendation to remove a legitimate application could be the correct result if it matches a policy. For example, a policy to refuse all Microsoft Office applications would recommend the removal of Microsoft Word. As long as the alert is clear that this is a policy decision and not a mistake then the product will not face a penalty.

For example, an acceptable alert would be: 'Word.exe is not permitted due to policy: NoMicrosoft', whereas an unacceptable alert would be: "Word.exe is a threat that should be removed (Trojan.XYZ)".

We think that measuring NOCAs is more useful than simply counting rarer FPs. The table below shows how we score different combinations of Classifications (the vertical axis) and Actions (the horizontal axis).

| | Recommendation: None | Recommendation: Allow | Recommendation: Unclear | Recommendation: Remove | Action: Remove |
|---|---|---|---|---|---|
| Safe | 2 | 1.5 | 1 | | |
| Unknown | 2 | 1 | 0.5 | 0 | -0.5 |
| Not Classified | 2 | 0.5 | 0 | -0.5 | -1 |
| Suspicious | 0.5 | 0 | -0.5 | -1 | -1.5 |
| Unwanted | 0 | -0.5 | 1 | -1.5 | -2 |
| Malicious | | | | 2 | -2 |

### Legitimate Software Prevalence Rating Modifiers

| | |
|---|---|
| Very High Impact | 5 |
| High Impact | 4 |
| Medium Impact | 3 |
| Low Impact | 2 |
| Very Low Impact | 1 |

### Prevalence Ratings

There is a significant difference between a product incorrectly alerting against a popular application like Microsoft Word and condemning a rare, obscure or

outdated application such as Internet Explorer 6. One is very popular all over the world and its detection as malware (or something less serious, but still suspicious) is a big deal.

Conversely, the outdated web browser has not been in general use for years and in many cases should not be used in a business environment. Detecting this application as malware may be wrong (an FP) but the mistake is less impactful.

With this mind, we collected objects of varying popularity and sorted them into five separate categories, as follows:

**1. Very High Impact**
**2. High Impact**
**3. Medium Impact**
**4. Low Impact**
**5. Very Low Impact**

Incorrectly labelling any legitimate object invokes penalties, but classifying Microsoft Word as malware, and recommending its removal without providing any context, will bring far greater penalties than doing the same for an ancient, unsupported web browser.

In order to calculate these relative penalties, we assign each impact category with a rating modifier, as shown in the table above.

Objects are obtained from original sources in most cases, avoiding third-party download sites. This is due to the risk of third parties modifying the legitimate objects and potentially adding problematic elements that could be a threat to an organisation. We remove adware and other less obviously legitimate objects from the test set.

We base the prevalence for each object on publicly available data sources.

## Accuracy Ratings

We calculate legitimate interaction ratings by multiplying together the interaction and prevalence ratings for each object:

**Accuracy Rating = Interaction Rating x Prevalence Rating**

## Legitimate Interaction Rating

| Product | None (allowed) | Click to Block | None (blocked) |
|---|---|---|---|
| AhnLab EPP/EDR | 100 | 0 | 0 |

- **Products that do not bother users and classify most applications correctly earn more points than those that ask questions and condemn legitimate applications.**

If a product inspected one legitimate, Medium Impact application and gave no alert or recommendation, its Accuracy Rating would be calculated like this:

**Accuracy Rating = 2 x 3 = 6**

If it labelled the object as 'suspicious' its rating would be calculated like this:

**Accuracy Rating = 0.5 x 3 = 1.5**

This same calculation is made for each legitimate object in the test and the results are summed and used to populate the graph and table shown under **5. Legitimate Accuracy Rating** in this report.

## Distribution of Impact Categories

In this test there was a range of objects with different levels of prevalence. The table below shows the frequencies:

## Legitimate Software Category Frequency

| Prevalence Rating | Frequency |
|---|---|
| Very High Impact | 0 |
| High Impact | 41 |
| Medium Impact | 38 |
| Low Impact | 21 |
| Very Low Impact | 0 |

# Appendix B: Terms Used

**Compromised** The attack succeeded, resulting in malware running unhindered on the target. In the case of a targeted attack, the attacker was able to take remote control of the system and carry out a variety of tasks without hindrance.

**Blocked** The attack was prevented from making any changes to the target.

**False Positive** When a security product misclassifies a legitimate application or website as being malicious, it generates a 'false positive'.

**Neutralised** The exploit or malware payload ran on the target but was subsequently removed.

**Complete Remediation** If a security product removes all significant traces of an attack, it has achieved complete remediation.

**Target** The test system that is protected by a security product.

**Threat** A program or sequence of interactions with the target that is designed to take some level of unauthorised control of that target.

**Update** Security vendors provide information to their products in an effort to keep abreast of the latest threats. These updates may be downloaded in bulk as one or more files or requested individually and live over the internet.

# Appendix C: FAQs

**Q What is a partner organisation? Can I become one to gain access to the threat data used in your tests?**

A Partner organisations benefit from our consultancy services after a test has been run. Partners may gain access to low-level data that can be useful in product improvement initiatives and have permission to use award logos, where appropriate, for marketing purposes. We do not share data on one partner with other partners. We do not partner with organisations that do not engage in our testing.

**Q We are a customer considering buying or changing our endpoint protection and/ or endpoint detection and response (EDR) product. Can you help?**

A Yes, we frequently run private testing for organisations that are considering changing their security products.
Please contact us at **info@selabs.uk** for more information.

## A **full methodology** for this test is available from our website.

- The test was conducted between 25th September and 3rd October 2025.
- All products were configured according to each vendor's recommendations, when such recommendations were provided.
- Targeted attacks were selected and verified by SE Labs.
- Malicious emails, URLs, attachments and legitimate messages were independently located and verified by SE Labs.
- Malicious and legitimate data was provided to partner organisations once the test was complete.

# Appendix D: Attack Details

## Wizard Spider

| Incident no: | Delivery | Execution | Action | Privilege Escalation | Post-Esclation Action | Lateral Movement | Lateral Action |
|---|---|---|---|---|---|---|---|
| 1 | Spear Phishing Attachment | Windows Command Shell | File and Directory Discovery | Bypass User Account Control | Remote System Discovery | Service Execution | Archive Collected Data |
| | | Malicious File | Process Discovery | Valid Accounts | Security Software Discovery | Domain Accounts | Data staged |
| | | Obfuscated Files or Information | System Information Discovery | | LLMNR/NBT-NS Poisoning and SMB Relay | | Data from Local System |
| | | Powershell | System Network Configuration Discovery | | | | Exfiltration Over C2 Channel |
| | | | System Owner/User Discovery | | | | |
| 2 | Spear Phishing Link | Malicious Link | File and Directory Discovery | Bypass User Account Control | NTDS | SSH | Archive Collected Data |
| | | Windows Command Shell | Process Discovery | Valid Accounts | Security Account Manager | External Remote Services | Data staged |
| | | Web Protocols | System Information Discovery | | Kerberoasting | | Data from Local System |
| | | Non-standard Port | Permission Groups Discovery | | | | Exfiltration Over C2 Channel |
| | | | System Owner/User Discovery | | | | |
| 3 | Spear Phishing Attachment | Malicious File | File and Directory Discovery | Bypass User Account Control | Windows Service | Lateral Tool Transfer | Archive Collected Data |
| | | Windows Command Shell | Process Discovery | Valid Accounts | Registry Run Keys / Startup Folder | Remote Desktop Protocol | Data staged |
| | | Web Protocols | System Information Discovery | | Scheduled Task | SMB/Windows Admin Shares | Data from Local System |
| | | | System Owner/User Discovery | | Masquerade Task or Service | | Exfiltration Over C2 Channel |
| | | | | | Winlogon Helper DLL | | |
| 4 | Spear Phishing Link | Malicious Link | File and Directory Discovery | Bypass User Account Control | Dynamic-link Library Injection | Windows Remote Management | Archive Collected Data |
| | | Windows Command Shell | Process Discovery | Valid Accounts | Windows File and Directory Permissions Discovery | | Data from Local System |
| | | Web Protocols | System Information Discovery | | | | Exfiltration Over C2 Channel |
| | | | System Network Configuration Discovery | | | | |

# Appendix E: Product Version

The table below shows the service's name as it was being marketed at the time of the test.

| Vendor | Product | Build Version (start) | Build Version (end) |
|---|---|---|---|
| AhnLab | AhnLab EPP/EDR | **Server Version:**<br>AhnLab EPP Management: 1.0.19.29 (Build 10069)<br>AhnLab EDR Analyzer: 2.0.9.8 (Build 3464) | **Server Version:**<br>AhnLab EPP Management: 1.0.19.29 (Build 10069)<br>AhnLab EDR Analyzer: 2.0.9.8 (Build 3464) |
| | | **DC:**<br>AhnLab Security Agent (EPP) for Windows: 1.0.19.14 (Build 2000)<br>AhnLab EDR Agent for Windows: 2.0.9.9 (Build 645)<br>AhnLab V3 Net for Windows Server: 9.0.90.7 (Build 2135) | **DC:**<br>AhnLab Security Agent (EPP) for Windows: 1.0.19.14 (Build 2000)<br>AhnLab EDR Agent for Windows: 2.0.9.9 (Build 645)<br>AhnLab V3 Net for Windows Server: 9.0.90.7 (Build 2135) |
| | | **PC:**<br>AhnLab Security Agent (EPP) for Windows: 1.0.19.14 (Build 2000)<br>AhnLab EDR Agent for Windows: 2.0.9.9 (Build 645)<br>AhnLab V3 Endpoint Security: 9.0.90.7 (Build 2135) | **PC:**<br>AhnLab Security Agent (EPP) for Windows: 1.0.19.14 (Build 2000)<br>AhnLab EDR Agent for Windows: 2.0.9.9 (Build 645)<br>AhnLab V3 Endpoint Security: 9.0.90.7 (Build 2135) |
| | | **Linux:**<br>AhnLab Security Agent (EPP) for Linux: 1.0.19.11 (Build 97)<br>AhnLab EDR Agent for Linux: 2.0.9.2000 (Build 5)<br>AhnLab V3 Net for Linux Server: 3.6.21.5 (Build 1373) | **Linux:**<br>AhnLab Security Agent (EPP) for Linux: 1.0.19.11 (Build 97)<br>AhnLab EDR Agent for Linux: 2.0.9.2000 (Build 5)<br>AhnLab V3 Net for Linux Server: 3.6.21.5 (Build 1373) |

SE LABS

## SE Labs Report Disclaimer

1. The information contained in this report is subject to change and revision by SE Labs without notice.
2. SE Labs is under no obligation to update this report at any time.
3. SE Labs believes that the information contained within this report is accurate and reliable at the time of its publication, which can be found at the bottom of the contents page, but SE Labs does not guarantee this in any way.
4. All use of and any reliance on this report, or any information contained within this report, is solely at your own risk. SE Labs shall not be liable or responsible for any loss of profit (whether incurred directly or indirectly), any loss of goodwill or business reputation, any loss of data suffered, pure economic loss, cost of procurement of substitute goods or services, or other intangible loss, or any indirect, incidental, special or consequential loss, costs, damages, charges or expenses or exemplary damages arising his report in any way whatsoever.
5. The contents of this report does not constitute a recommendation, guarantee, endorsement or otherwise of any of the products listed, mentioned or tested.
6. The testing and subsequent results do not guarantee that there are no errors in the products, or that you will achieve the same or similar results. SE Labs does not guarantee in any way that the products will meet your expectations, requirements, specifications or needs.
7. Any trade marks, trade names, logos or images used in this report are the trade marks, trade names, logos or images of their respective owners.
8. The contents of this report are provided on an “AS IS” basis and accordingly SE Labs does not make any express or implied warranty or representation concerning its accuracy or completeness.