# Measuring the health of antivirus ecosystems

Fanny Lalonde Lévesque
Ecole Polytechnique de Montréal
Montréal, Canada
fanny.lalonde-levesque@polymtl.ca

Anil Somayaji
Carleton University
Ottawa, Canada
soma@scs.carleton.ca

Dennis Batchelder
Microsoft Corporation
Redmond, United States
dennis.batchelder@microsoft.com

José M. Fernandez
Ecole Polytechnique de Montréal
Montréal, Canada
jose.fernandez@polymtl.ca

## Abstract

*The number and variety of computer threats has fueled a digital arms race, resulting in a complex software ecosystem around malware and antivirus (anti-malware) products. While there has been significant past work in benchmarking antivirus (AV) products against each other, how healthy is the overall AV software ecosystem?*

*Using data collected from Microsoft Windows Malicious Software Removal Tool (MSRT) running on more than one billion machines, we develop ecosystem health measures based upon infection rates, product diversity, market dominance, and activity status. Our study shows that while a diverse group of products is used and the vast majority of them are running properly, there is also significant churn in product usage which may indicate dissatisfaction with current products. While further work is needed to better understand these patterns, this study shows the potential power of an ecosystem health-based approach to studying AV performance in practice.*

## 1 Introduction

The multitude of security products available on the market has evolved into a complex ecosystem that interacts with the malware landscape. Given the increase in complexity and diversity of both malware threats and antivirus (AV) products, the evaluation of the latter is essential in helping the industry develop better products that match the evolving nature of malware and meet users' expectation.

While typical evaluation methods are mostly focused on single-product or comparative tests, AV products are always evaluated on an individual basis and not as a whole. Measuring the overall performance of AV products can provide a better understanding of their global condition and help identify issues that could not be studied using current AV testing methods. Moreover, such evaluation can allow the investigation of the aggregated effect, if any, of AV products beyond their individual contribution.

One approach is to consider AV products as software ecosystems, that is as a collection of software solutions that are developed and co-evolve in the same environment. The concept of *ecosystem health*, which refers to the global condition of an ecosystem, provides a powerful theoretical and practical framework for monitoring system activity, identifying and predicting areas for improvement, and evaluating changes in ecosystems [2]. Applied to AV products, the "health" of AV ecosystems can be measured as its overall performance, that is how well it is protecting users against specific, prevalent malware. Developing relevant indicators such as the number and relative usage of different AV products, or how well maintained those installations are, could allow to track the status of the AV ecosystem and assess its overall condition and quality.

In this paper we report on the first study that aimed to define and measure the health of AV ecosystems by developing scalable indicators in terms of activity, diversity and stability. Using four months of sampled telemetry data from Microsoft Software Removal Tool (MSRT) on millions of computers, we analyse AV product status (whether they are running and have up to date signatures), AV vendor diversity, and AV stability in terms of protection status and security vendor. We examine some initial testing to investigate how those indicators relate to MSRT infection rates (of malware missed by the installed AV), discuss opportunities for future testing and successfully identify areas that could be improved within the AV ecosystem.

The remainder of the paper is organized as follows. Section 2 presents the literature and related work in natural and

978-1-5090-0319-8/15/$31.00 ©2015 IEEE

software ecosystems health. In Section 3 we describe our indicators of AV ecosystems and discuss their associations with users' protection in Section 4. We discuss methodological limitations and future work in Section 5 and conclude in Section 6.

## 2 Background

First introduced in the field of natural ecosystems, Costanza [4] defines a healthy ecosystem as being "stable and sustainable". While there is no universally accepted definition and indicators, ecosystem health could be defined as a combined measure of system vigor (productivity), organization (including diversity and interactions) and resilience. Vigor or productivity refers to the capacity of the system to sustain its activity. Organization refers to number and diversity of interactions between components of the system. Resilience refers to the ability of the system to maintain its structure and activity in the presence of stress. That is, a healthy natural ecosystem is one that can develop an efficient diversity of components and exchange pathways (organization) while maintaining its activity (vigor) over time in the face of stress (resilience).

Beyond natural ecosystems, the notion of ecosystem health has also inspired the field of software ecosystem (SECO), where health refers to how well the ecosystem is functioning, that is its ability to endure and remain variable and productive over time [12]. For example, Wynn [19] applied the concept of natural ecosystem health to develop a framework in terms of vigor, resilience and organization to gauge the health and sustainability of open source projects.

SECO health has also been applied with a business ecosystem (BECO) approach, as a means of expanding development, better positioning in the market, or increasing revenues. In the BECOs health literature, the concept of health is defined as the ability of a BECO to provide "durably growing opportunities for its members and those who depend on it" [7]. The notions of vigor, organization and resilience are adopted and changed to productivity, niche creation and robustness [7, 8, 9].

While ecosystem health has been applied to various SECOs, such as open source software [10, 16] or hardware-dependent software [18], it has not been used significantly in the area of AV software. Many have talked about ecosystems and ecosystem-related concepts in computer security, particularly with regards to monocultures [6] and mechanisms for automated software diversity [5]. When assessing the performance of AV systems, however, the main focus has traditionally been on single-product or comparative tests of AV systems' ability to detect malware and ignore benign software. Whether performed in a controlled lab environment or through field studies [15, 11], current evaluation methods are limited to the individual performance of security products.

Security vendors have also used software telemetry data for quality assurance. However, those analysis are not intended to study the overall performance of AV systems, but rather focus on one single vendor. Closer to our research is the work done by Blackbird *et al.* [3], where they used MSRT data to evaluate the global impact of anti-malware protection state on infection rates of protected users. To the best of the authors' knowledge, there has been no previous work published in the literature based on such telemetry data to assess the overall health of AV ecosystems. The key contributions of this work are therefore the proposing of measures of AV ecosystem health and assessing that health using large-scale security software telemetry data.

## 3 Antivirus ecosystem indicators

The evaluation of AV ecosystems requires the development of relevant, scalable, easy to measure and understand indicators. In this work we propose to characterize AV ecosystems in terms of activity, diversity, and stability. Using our indicators, we conduct a longitudinal analysis to track the status of the global AV ecosystem over a 4-month period.

The data was collected by MSRT, a malware cleaner utility that scans computers for infections of specific, prevalent malicious software and helps remove these infections. MSRT is delivered and runs every month on more than one billion machines through Windows Update as well as being available as a separate download from Microsoft. It is worth mentioning that MSRT only runs on Windows and that it, by design, only detects a subset of the malware families covered by Windows Defender and other Microsoft anti-malware products. Upon its execution, MSRT also calls the Windows Security Center (WSC) API to collect information about the protection state of computers, such as the AV actively protecting the machine and its signature status. The data used in this analysis was collected from June 2014 to September 2014 on computers running Windows XP, Vista, 7, 8 and 8.1. As not all Windows users send their data to MSRT, we randomly selected one on every ten unique computers in order to limit potential effect of self-selection bias, reducing our sample size from one billion to 100+ million computers. Moreover, we restricted our analysis to the approximately 90% of computers that had an AV product installed, giving us a sample population of 90+ million hosts.

### 3.1 Activity

We define the activity of the AV ecosystem as the percentage of users with at least one AV product actively running with up to date signatures. Figure 1 illustrates the evolution of the percentage of users having an up to date AV

installed. Over the studied period, the activity of the AV ecosystem ranged from 87.50% to 88.60%.

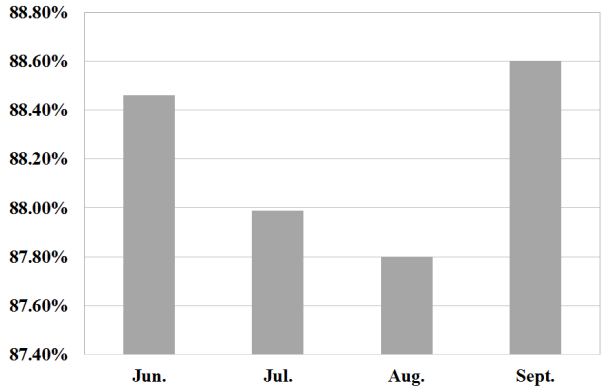


**Figure 1. AV activity over the 4 months**

More details on the status of AV products during the 4 months are presented in Table 1. The status enabled refers to an actively running AV product using the latest signature files available while out of date refers to an actively running AV product using out of date signatures. Expired refers to an actively running expired AV product. Snoozed means that the AV product is active but is not performing real-time monitoring, typically because the product is upgrading itself, and off means that the AV product is not running, as it has been turned off. Table 1 shows that within the users that have an AV product installed, between 11.40% and 12.21% are not protected with up to date signatures, despite having an AV installed.

**Table 1. AV status over the 4 months**

| | Jun. | Jul. | Aug. | Sept. |
|---|---|---|---|---|
| Enabled | 88.46% | 87.99% | 87.80% | 88.60% |
| Out of date | 3.74% | 3.91% | 3.96% | 3.74% |
| Expired | 1.87% | 2.23% | 2.22% | 1.98% |
| Snoozed | 0.45% | 0.43% | 0.33% | 0.30% |
| Off | 5.48% | 5.44% | 5.70% | 5.38% |

## 3.2 Diversity

The diversity of the AV ecosystem was characterized based on its richness, degree of concentration, and dominance.

In natural ecosystems, the richness (S) refers to the total number of different species. Applied to the global AV ecosystem, the richness can be defined as the total number of AV vendors within the ecosystem. Table 2 shows the evolution of the different indicators in terms of diversity. Over the 4 months, the richness did not vary from 107, indicating that the number of AV vendors was constant.

The degree of concentration (D), also known as Simpson's diversity Index [14] or Gini-Simpson Index, is a measure of the degree of concentration of individuals classified into types. It can be interpreted as the probability that two organisms belong to different species. The value S represents the richness –the total number of different AV vendors– and $p_i$ represents the fraction of AV products that belong to the $i$th AV vendor. A value of 0 indicates no diversity and 1.0 indicates high diversity.

$$D = 1 - \sum_{i=1}^{S} p_i^2$$

Based on the results in Table 2, we can tell that the AV ecosystem is highly diversified, as its degree of concentration varies around 0.92.

Dominance of the AV ecosystem was measured using the Berger-Parker Index (BP) [1]. This index estimates dominance using the prevalence of the most abundant type, which refers to the AV vendor with the highest market share. Results in Table 2 show that the dominance of the AV ecosystem varies between 0.136 and 0.143. In economics, that would indicate that the AV market has a low concentration (<0.5), ranging from perfect competition to an oligopoly.

**Table 2. AV diversity over the 4 months**

| | Jun. | Jul. | Aug. | Sept. |
|---|---|---|---|---|
| Richness | 107 | 107 | 107 | 107 |
| Concentration | 0.906 | 0.906 | 0.906 | 0.928 |
| Dominance | 0.176 | 0.179 | 0.136 | 0.181 |

## 3.3 Stability

Stability of the AV ecosystem was evaluated in terms of changes in AV status and AV vendor. Table 3 shows for each month the percentage of users by status that have a different AV status compared to the previous month. For example, the value in the first line (enabled) under the column Jul. means that 3.19% of the users that had an enabled AV product for June had a different AV status for July. The next value in the same line means that 3.57% of the users that had an enabled AV product for July had a different status for August. Interestingly, the rate of changes are not equivalent between the different AV status. Users with snoozed AV products, followed by out of date AV products, are the status with the highest rate of changes. The monthly rate of AV state changes varies between 10.84% and 11.91% (Table 3) and overall, 40.47% of the users changed their AV status over the 4 months.

In order to better understand the nature of those changes, we analysed for each month the percentage of users that switched from one AV status to an other. We found that more than 75% of the variations for enabled AV products went to either an out of date or off AV status. For all the other statuses (out of date, expired, snoozed, off), 75% of all changes in AV status went to an enabled AV status.

**Table 3. AV state changes over the 4 months**

| | Jul. | Aug. | Sept. |
|---|---|---|---|
| Enabled | 3.19% | 3.57% | 3.16% |
| Out of date | 32.25% | 36.05% | 37.90% |
| Expired | 18.59% | 15.53% | 23.14% |
| Snoozed | 49.62% | 53.81% | 45.76% |
| Off | 19.85% | 20.51% | 25.92% |

We adopted a similar approach the evaluate the stability of the AV ecosystem in terms of changes in AV vendors. Overall, 33.57% of the users switched to a different AV vendor over the study (Figure 1).

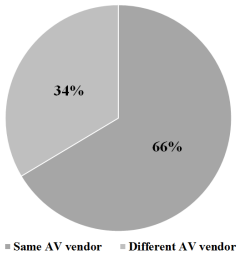


**Figure 2. Overall AV vendor changes**

We also investigated potential relationships between changes in AV status and AV vendors. As presented in Table 4, we can see that stability in AV vendors is associated with higher stability in AV status. To the opposite, 87.10% of the users that changed their AV vendor also experienced changes in the status of their AV product.

**Table 4. Overall AV stability**

| | Stable Status | Different Status |
|---|---|---|
| Stable AV vendor | 83.09% | 16.91% |
| Different AV vendor | 12.90% | 87.10% |

## 4 Country level analysis and evaluation

We define the health of the AV ecosystem as the measure of its aggregated performance, that is how well it is protecting users. As illustrated in Figure 3, we can see that the overall infection rates of users that had an AV product installed depend of the protection status of the latter. Not surprisingly, users with an enabled AV products have the lowest infection rates.

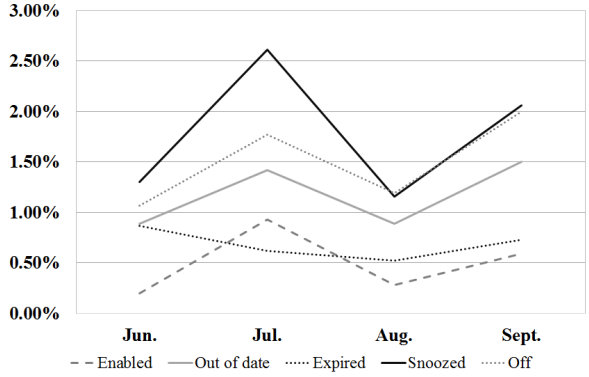


**Figure 3. Infection rates over the 4 months**

To evaluate how our indicators relate to users' protection in terms of infection rates, we conducted an empirical study of AV ecosystems defined by geographical unit, as classified by MSRT. While our data set is large overall, for some countries our sampled population is too small to allow for proper analysis. To determine the minimum representative sample size for each country, we performed a power analysis. We used a two-tailed one proportion Chi-Square test with a desired power of 90% and a level of significance of 1%. The minimum sample size computed was 37 149, which can be rounded to 38 000. We then excluded all countries that had less than 38 000 reports over the 4 months, reducing our sample from 187 to 126 countries.

Our sample of MSRT data contained many users that only report data sporadically. Because we are interested here in health changes over time, we limited our analysis to the portion of users that reported for all four months of the study period. This exclusion removed roughly two-thirds of the sample, leaving us with approximately 32+ million users.

Infection rates by country were used as the dependent variable in order to estimate the aggregated performance of AV ecosystems. The independent variables were selected to capture the activity, diversity and stability of AV ecosystems. The activity was represented by the proportion of users that had an actively running AV product with up to date signatures for the entire period of the study. The diversity was evaluated based on the 4-month averaged richness, degree of concentration, and dominance. The stability was computed by the proportion of users that experienced changes in AV status and the proportion of users that changed AV vendor during the studied period. Descriptive statistics for each factor are presented in Table 5. The mean allows to measure the central tendency of the data and the

standard deviation measures how concentrated the data are around the mean; the more concentrated, the smaller the standard deviation. From Table 5, we can see that richness varies widely across the different ecosystems. To the opposite, %Protected and concentration present small dispersion.

**Table 5. Descriptive statistics**

| Dimension | Indicator | Mean | Std. Dev. |
|---|---|---|---|
| Activity | %Protected | 0.849 | 0.049 |
| Diversity | Richness | 50.14 | 14.01 |
| | Concentration | 0.827 | 0.040 |
| | Dominance | 0.296 | 0.067 |
| Stability | %Unstable AV status | 0.482 | 0.091 |
| | %Unstable AV vendor | 0.399 | 0.085 |

In order to measure potential dependence between the dependent variable and our indicators, we first computed the Pearson correlation coefficients between the indicators and the infection rates (see Table 6). The value $r$ represents the correlation coefficient. A value of 1 implies that there is a linear relationship between the two factors; as one increases, the other also increases. To the opposite, a value of -1 means that when one value increases, the other decreases. And a value of 0 indicates that there is no linear correlation between the variables. The p-value was also computed to measure the significance of the results. A low p-value (such as 0.01) means that there is a 1 in 100 chance that we would have obtained the same results if the variables were not correlated. For the purpose of our analysis, we considered a correlation to be significant if the p-value was lower than 0.01. Results show that activity, diversity in terms of richness and stability are significantly correlated with the infection rate (p-value$<$0.001). Activity and richness are found to be negatively associated with the infection rate, meaning that high values are associated with low infection rates. To the opposite, higher changes in AV status or AV vendors are associated with higher infection rates.

**Table 6. Pearson correlation coefficients between infection rate and indicators (N=126 countries)**

| Indicator | $r$ | p-value |
|---|---|---|
| %Protected | -0.59 | 2.14e-13* |
| Richness | -0.42 | 6.99e-07* |
| Concentration | -0.08 | 3.70e-01 |
| Dominance | 0.15 | 8.51e-02 |
| %Unstable AV status | 0.71 | 7.03e-21* |
| %Unstable AV vendor | 0.67 | 8.16e-18* |

Although the Pearson correlation coefficient provides insight on the dependence between the infection rates and the indicators, its primary drawback is that it is very difficult to draw conclusions about the effect of one single factor on the dependent variable, as factors often interact together. We therefore conducted a multiple regression to examine the relative importance of each indicator. Multiple regression was selected as it allows to estimate the effect of the factors while controlling for the many factors that simultaneously affect the dependent variable. Because we are interested to assess the unique effect of each indicator, we looked for multicollinearity as it can reduce the effective amount of information available to evaluate the effect of the indicators. The presence of multicollinearity was investigated by computing the Pearson correlation coefficient matrix. Results in Table 8 show the presence of a very strong correlation ($r > 0.90$) between the two indicators related to stability. We therefore excluded the indicator %Unstable AV status from our analysis and only kept %Unstable AV vendor to estimate the stability of AV ecosystems.

**Table 7. Multiple general linear regression (N=126 countries)**

| Indicator | $\beta$ | Std. Error | t-value | p-value |
|---|---|---|---|---|
| %Protected | -0.33 | 0.08 | -4.18 | 5.50e-05* |
| Richness | -0.09 | 0.07 | -1.25 | 2.14e-01 |
| Concentration | 0.26 | 0.13 | 1.86 | 6.46e-02 |
| Dominance | 0.39 | 0.14 | 2.76 | 6.73e-03* |
| %Unstable AV vendor | 0.41 | 0.08 | 4.98 | 2.19e-06* |
| $R^2$ adjusted | 0.53 | | | |
| F-statistic | 29.25 | | | |
| Degree of freedom | 5 | | | |
| Df (residuals) | 120 | | | |
| p-value | 2.57e-19 | | | |

Table 7 presents the results from the multiple general linear regression. For each factor, the standardized regression coefficient $\beta$ and its associated standard error (Std. Error) were computed. The p-value, which is interpreted as an indicator of the significance of the results, was also computed: a low p-value indicates that the null hypothesis can be rejected with high confidence, and that the variable is relevant in the regression model. In order to limit potential effect of chance, that is to discover a significant correlation purely by chance, we considered a relationship to be significant if the p-value was lower than 0.01. We also provided the t-value of each factor, which provides insight on the direction (positive or negative) and magnitude of the effect. The results of the multiple regression (see Table 7) indicate that %Protected, dominance and %Unstable AV vendor have a statistically significant (p$<$ 0.01) relationship with the infection rate. While %Protected is found to have a negative association with the infection rate, dominance and %Unstable AV vendor have a positive relationship.

**Table 8. Pearson correlation coefficient matrix between indicators (N=126 countries)**

| | %Protected | Richness | Concentration | Dominance | %Unstable AV status | %Unstable AV vendor) |
|---|---|---|---|---|---|---|
| %Protected | 1.00 | 0.38*** | -0.08 | 0.10 | -0.69*** | -0.60*** |
| Richness | - | 1.00 | 0.05 | -0.05 | -0.05*** | -0.49*** |
| Concentration | - | - | 1.00 | -0.90*** | -0.05 | -0.05 |
| Dominance | - | - | - | 1.00 | 0.09 | 0.07 |
| %Unstable AV status | - | - | - | - | 1.00 | 0.98*** |
| %Unstable AV vendor | - | - | - | - | - | 1.00 |

*Statistically significant at 0.05 level; **Statistically significant at 0.01 level; ***Statistically significant at 0.001 level.

As the infection rates are function of the protection status (see Figure 3), we investigated to see if our previous findings apply to users when stratified by protection status. We classified users as being protected if they had an enabled AV during the entire study and unprotected if they had either an out of date, expired, snoozed or off AV, or no AV installed. Protected users got an average infection rate of 1.33% (SD=0.0067, 95% CI=0.0059-0.0076) while unprotected users got an average infection rate of 21.43% (SD=0.1315, 95% CI=0.1171-0.1501). As a comparison, the average infection rate for all users having an AV installed, regardless of the status of the latter, was 2.02% (SD=0.0124, 95% CI=0.0110-0.0141).

**Table 9. Pearson correlation coefficients between infection rates and indicators by protection status (N=126)**

| | $r$ | p-value |
|---|---|---|
| Protected | | |
| %Protected | -0.51 | 1.08e-09* |
| Richness | -0.40 | 2.63e-06* |
| Concentration | 0.01 | 9.73e-01 |
| Dominance | 0.05 | 5.71e-01 |
| %Unstable AV status | 0.60 | 6.54e-14* |
| %Unstable AV vendor | 0.59 | 4.29e-13* |
| Unprotected | | |
| %Protected | -060 | 9.97e-14* |
| Richness | -0.40 | 4.24e-06* |
| Concentration | -0.06 | 5.19e-01 |
| Dominance | 0.13 | 1.45e-01 |
| %Unstable AV status | 0.86 | 0.00e-01* |
| %Unstable AV vendor | 0.84 | 0.00e-01* |

The Pearson correlation coefficients were first computed to identify any potential statistical association between the indicators and the infection rates by protection status. From Table 9 we can see that the relationships do not differ between protected and unprotected users. Moreover, the correlations found are similar to our previous findings (see Table 6): high %Protected and richness are associated with lower infection rates and high instability is associated wih higher infection rates.

To better estimate the unique effect of each indicator, we performed a multiple general linear regression for each protection status (e.g. protected and unprotected). It appears the main difference between protected and unprotected users is the effect of dominance (see Table 10). While dominance is not related to infection rates for protected users, a negative significant association is found for unprotected users, meaning that higher dominance is associated with higher infection rates for unprotected users but not for protected.

## 4.1 Activity

The country level analysis allowed to identify a statistically significant correlation between the proportion of computers running an enabled AV product and malware infection rates. Intuitively, as the proportion of users running an enabled AV product increases, the rate of malware infections among users that have an AV product installed decreases. However, what is less intuitive, is that the infection rates for unprotected users also tend to be lower in countries with higher proportion of users protected.

A first explanation for this is that unprotected users benefit from a herd immunity effect from protected users. This explanation can be explored by examining the correlation between the protection coverage –the proportion of protected users among all users– and the infection rates for unprotected users. A strong positive relationship ($r$=-0.85, p-value=0.00e-01, N=127) was found between the protection coverage and the infection rates for unprotected. As shown in Figure 4, higher protection coverage is associated with lower infection rates for unprotected users. Although this broad correlation may provide empirical evidence of a herd immunity effect, proper validation should be achieved by conducting further studies designed for the purpose.

A second possibility is that the proportion of users running an enabled AV product acts as a country level marker for users investment in the security of their computers. This would indicate that users in countries with higher protection coverage are less infected because they tend to be more aware of security risks and less likely to engage in risky behaviours. The validation of this explanation would re-

**Table 10. Multiple general linear regression by protection status (N=126 countries)**

| | Protected | | | | Unprotected | | | |
|---|---|---|---|---|---|---|---|---|
| Indicator | $\beta$ | Std. Error | t-value | p-value | $\beta$ | Std. Error | t-value | p-value |
| %Protected | -0.29 | 0.07 | -4.04 | 9.62e-05* | -0.19 | 0.06 | -3.13 | 2.19e-03* |
| Richness | 0.01 | 0.06 | 0.24 | 8.13e-01 | 0.04 | 0.05 | 0.69 | 4.91e-01 |
| Concentration | 0.22 | 0.12 | 1.82 | 7.03e-02 | 0.22 | 0.11 | 2.14 | 3.43e-02 |
| Dominance | 0.31 | 0.12 | 2.52 | 1.32e-02 | 0.30 | 0.11 | 2.82 | 5.53e-03* |
| %Unstable AV vendor | 0.58 | 0.07 | 7.82 | 2.34e-12* | 0.73 | 0.06 | 11.38 | 0.00e-01* |
| | $R^2$ adjusted | | | 0.62 | $R^2$ adjusted | | | 0.72 |
| | F-statistic | | | 42.20 | F-statistic | | | 65.85 |
| | Df | | | 5 | Df | | | 5 |
| | Df (residuals) | | | 120 | Df (residuals) | | | 120 |
| | p-value | | | 0.00e-01 | p-value | | | 0.00e-01 |

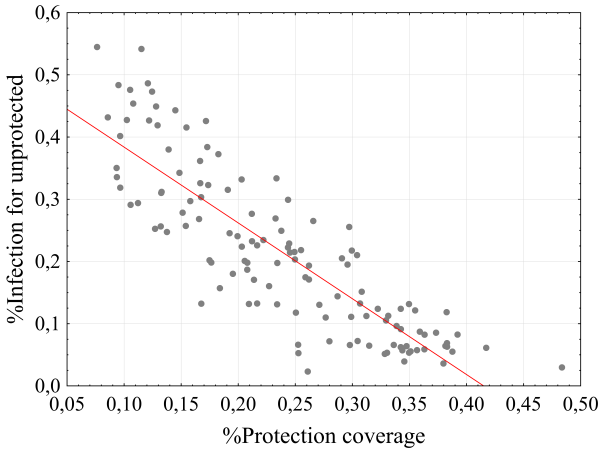


**Figure 4. Infection rates for unprotected function of the protection coverage**

quire to conduct either country level studies based on aggregated measures of user security awareness or large-scale user studies.

**Findings:** Among users that have an AV product installed, 10% are not actively protected with up to date signatures.

Higher AV activity is significantly associated with lower infection rates, regardless of the protection status.

## 4.2 Diversity

Measures of diversity in terms of richness and dominance were significantly correlated with the infection rates. The degree of concentration, however, was not found to be significant in either of the regression models.

Richness was significantly and negatively correlated with infection rates of all users, whether protected or not. However, the correlation was not significant in the regression model after controlling for the other factors. One explanation could be that countries with higher AV adoption are more likely to have a diversified AV market.

A positive significant correlation was found between dominance and infection rates based on the regression models for all users and unprotected users, but not for protected users. One potential explanation could be that users in countries with AV monoculture are more vulnerable. The term monoculture refers originally to an agricultural practice of producing or growing one single crop over a broad area for several consecutive years. Since all plants are genetically similar, they are more vulnerable and less resistant to infections by pathogen, insects, or environmental conditions. If a new disease strikes to which they have no resistance, the entire population of crops can be destroyed. If we extend the principle to AV ecosystems, an AV monoculture would occur when the AV market is dominated by one single vendor. Therefore, population in such ecosystem are more likely to be infected when exposed to a new malware against which the AV product is not able to protect. This hypothesis should, however, be confirmed with theoretical models and validated with proper experiments in order to better understand the nature of the statistical association.

**Findings:** The global AV ecosystem has a degree of concentration around 0.92, meaning that it is highly diversified.

Higher AV dominance is significantly correlated with higher infection rates for all users that have an AV installed, as well as for unprotected users.

## 4.3 Stability

All indicators of AV stability were negatively significantly correlated with infection rates. These statistical associations were also found to be significant from the regression models for all users, as well as for protected and unprotected users. Moreover, the changes in AV vendors appeared to be the indicator with the strongest effect (based on its t-value) on infection rates.

One plausible explanation is that users switched vendors because they got infected. To examine this explanation, we

computed the rate of changes in AV vendors for users that did not get any infection over the 4 months. Overall, 8.20% of those cleaned users changed AV vendor, while the rate for all users was 33.57%. Changes in AV vendors could therefore be a consequence of detections by MSRT and be interpreted as a potential marker of users' satisfaction regarding the protection provided by AV systems. From those findings, AV vendors should make sure to detect the malware families covered by MSRT if they want to retain their customers.

**Findings:** Over the studied period, 40.47% of users changed their AV status and 33.57% changed AV vendor.

Higher stability, both in AV status and AV vendor, is significantly associated with lower infection rates for all users, whether having an AV product installed or not.

## 5 Discussion

This study and its results are subject to a number of limitations. First, there is an inherent bias to our results because our sample population is drawn from Windows systems running MSRT; thus, it does not provide insight into Windows systems that do not run Windows Update, and it does not give insight into the performance of AV on non-Windows systems such as MacOS and Android. However, given that there are more than one billion computers regularly running MSRT, patterns discovered in this population are important on their own, whether or not they are representative of patterns in other computational contexts.

Another significant limitation is that the infection rates as determined by MSRT are only for a subset of malware families. While these families may represent some of the most significant malware families on Windows, they are not a representative sample and so MSRT reported infection rates will be different from the overall malware infection rate. Nevertheless, given the significance of MSRT-targeted malware these infection rates are also of inherent interest.

This study was intended to be exploratory and not confirmatory, as our purpose was to develop indicators and investigate how they may relate to users' protection. Although ecosystem health measures cannot give predictive descriptions or identify causal mechanisms, they do provide case-by-case evaluations in real-world settings [17]. Further studies should be conducted in order to validate our findings and investigate the nature of the associations we found.

Our results regarding the diversity of AV products used in practice roughly agree with other assessments of product market share [13]. Perhaps the best news coming out of this study is that almost 90% of the observed Windows systems are protected by anti-malware software and almost 90% of those systems are actively scanning with up to date signatures. While these numbers mean that the vast majority of users are maintaining their systems properly, given the large number of Windows installations, these numbers also mean that millions of systems are not adequately protected.

There are also some clear indicators of ecosystem-wide dysfunction. Approximately one third of the systems running an AV product at the start of the study were using a different AV by the end of the four months. That is a remarkably high level of user churn; further, this churn is broadly distributed given the richness, high concentration, and low dominance of AV products in our sample. One hypothesis for this churn is that users are unsatisfied with AV products in general; clearly, though, this hypothesis requires further evaluation.

The country level analysis suggests the potential value of diversity in AV products, with countries with higher AV dominance having higher infection rates. The evidence that higher protection rates are correlated with lower infection rates in unprotected computers, suggests that mechanisms beyond the actual protection provided by AV products have protective effects, or that AV products provide protection for more than the host they are installed upon. Differentiating between these two effects may be a interesting area for future research as this question provides insight into how systems are or are not compromised by malware.

As AV products are continuously evolving, the process of evaluating the health of AV ecosystems should also evolve. In particular, while we hypothesize there is inherent value to diversity in the AV ecosystem and our work provides support for this hypothesis, it may be worth developing non-diversity based measures of ecosystem health in order to capture other important AV ecosystem-related health patterns. Further, a business ecosystem approach could be applied to evaluate AV ecosystem health defined by AV vendor, rather than by geographical unit. Such analysis could be seen as a complementary AV test to help customers choose an AV vendor based on specific indicators like users' loyalty, growth, or market share. Having said that, we believe the insights reported here show the potential benefits of our ecosystem health approach to studying AV performance.

## 6 Conclusion

In this paper we present a definition of antivirus ecosystem health based on a population's characteristic levels of activity, diversity, and stability. Using four months of telemetry data from MSRT, we calculated these health measures for a sample of more than one billion MSRT users and correlated them with MSRT-reported infection rates, in aggregate and on a per-country basis. Lowered infection rates were positively correlated with higher rates of AV activity, stable AV product usage and status, and AV product diversity. Higher AV activity also seems to be positively correlated with lowered infection rates on systems not pro-

tected by AV software. While the results of this study cannot be considered definitive, they suggest that further work into measures of AV ecosystem health may produce significant insights into the performance of antivirus systems in practice.

## 7 Acknowledgments

The authors would like to thank the Microsoft Malware Protection Center (MMPC) for granting us access to the MSRT telemetry data and for supporting this work.

## References

[1] W. H. Berger and F. L. Parker. Diversity of planktonic foraminifera in deep-sea sediments. *Science*, 168(3937):1345–1347, 1970.

[2] P. Bertollo. Assessing ecosystem health in governed landscapes: a framework for developing core indicators. *Ecosystem health*, 4(1):33–51, 1998.

[3] J. Blackbird and B. Pfeifer. The global impact of anti-malware protection state on infection rates. In *Virus Bulletin Conf.*, 2013.

[4] R. Costanza, B. G. Norton, and B. D. Haskell. *Ecosystem health: new goals for environmental management*. Island Press, 1992.

[5] S. Forrest, A. Somayaji, and D. H. Ackley. Building diverse computer systems. In *The Sixth Workshop on Hot Topics in Operating Systems*, pages 67–72. IEEE, 1997.

[6] D. Geer, R. Bace, P. Gutmann, P. Metzger, C. P. Pfleeger, J. S. Quarterman, and B. Schneier. Cyberinsecurity: The cost of monopoly. *How the dominance of Microsofts products poses a risk to society*, 2003.

[7] M. Iansiti and R. Levien. Keystones and dominators: Framing operating and technology strategy in a business ecosystem. *Harvard Business School, Boston*, 2004.

[8] M. Iansiti and R. Levien. Strategy as ecology. *Harvard business review*, 82(3):68–81, 2004.

[9] M. Iansiti and G. L. Richards. Information technology ecosystem: Structure, health, and performance, the. *Antitrust Bull.*, 51:77, 2006.

[10] S. Jansen. Measuring the health of open source software ecosystems: Beyond the scope of project health. *Information and Software Technology*, 56(11):1508–1519, 2014.

[11] F. Lalonde Levesque, J. Nsiempba, J. M. Fernandez, S. Chiasson, and A. Somayaji. A clinical study of risk factors related to malware infections. In *ACM SIGSAC Conf. on Computer & Communications Security*, pages 97–108. ACM, 2013.

[12] K. Manikas and K. M. Hansen. Reviewing the health of software ecosystems-a conceptual framework proposal. In *IWSECO@ ICSOB*, pages 33–44, 2013.

[13] OPSWAT. Antivirus and Threat Report: January 2014. `https://www.opswat.com/resources/reports/antivirus-january-2014`, 2014.

[14] E. H. Simpson. Measurement of diversity. *Nature*, 1949.

[15] A. Somayaji, Y. Li, H. Inoue, J. Fernandez, and R. Ford. Evaluating security products with clinical trials. In *USENIX Workshop on Cyber Security Experimentation and Test (CSET)*, 2009.

[16] S. Van Lingen, A. Palomba, and G. Lucassen. On the software ecosystem health of open source content management systems. In *5th International Workshop on Software Ecosystems (IWSECO 2013)*, page 38, 2013.

[17] B. A. Wilcox. Ecosystem health in practice: emerging areas of application in environment and human health. *Ecosystem Health*, 7(4):317–325, 2001.

[18] K. Wnuk, K. Manikas, P. Runeson, M. Lantz, O. Weijden, and H. Munir. Evaluating the governance model of hardware-dependent software ecosystems–a case study of the axis ecosystem. In *Software Business. Towards Continuous Value Delivery*, pages 212–226. Springer, 2014.

[19] D. Wynn Jr, M. Boudreau, and R. Watson. Assessing the health of an open source ecosystem. *Emerging Free and Open Source Software Practices*, pages 238–258, 2007.