

Decision making model design for antivirus software selection using Factor Analysis and Analytical Hierarchy Process

Ai Nurhayati^{1*}, Aditya Gautama², and Muchammad Naseer³

^{1,3} Sekolah Tinggi Teknologi Bandung, Bandung, Indonesia

² Politeknik Negeri Batam, Batam, Indonesia

Abstract. Virus spread increase significantly through the internet in 2017. One of the protection method is using antivirus software. The wide variety of antivirus software in the market tends to creating confusion among consumer. Selecting the right antivirus according to their needs has become difficult. This is the reason we conduct our research. We formulate a decision making model for antivirus software consumer. The model is constructed by using factor analysis and AHP method. First we spread questionnaires to consumer, then from those questionnaires we identified 16 variables that needs to be considered on selecting antivirus software. This 16 variables then divided into 5 factors by using factor analysis method in SPSS software. These five factors are security, performance, internal, time and capacity. To rank those factors we spread questionnaires to 6 IT expert then the data is analyzed using AHP method. The result is that performance factors gained the highest rank from all of the other factors. Thus, consumer can select antivirus software by judging the variables in the performance factors. Those variables are software loading speed, user friendly, no excessive memory use, thorough scanning, and scanning virus fast and accurately.

1 Introduction

Nowadays, almost every person is connected to the internet. It has a lot of impact in daily lives that can be positive or negative. One of the negative impact is the infection of computer virus, spreading faster through the internet.

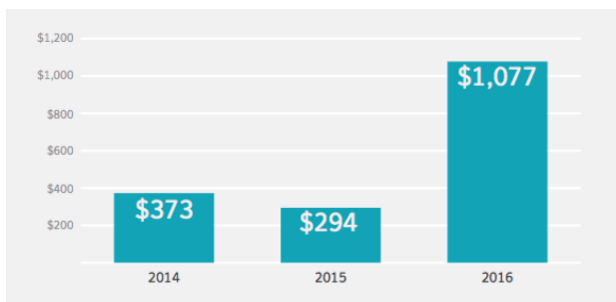


Fig 1. Internet Security Threat Report. (Source: Symantec 2017 Internet Security Threat Report)

This problem have made the amount of antivirus software increased in 2017. People who usually connected to the internet, such as students, businessman, and teacher, need to anticipate virus infection by using one of the antivirus software that are available in the market. The variety of antivirus software brand have made it difficult for them to choose which one is the best for their computer system. In this research, we will identify which variables that needs to be considered in choosing antivirus software. The methods that we used are through survey, questioner

and direct interview with the user. The variables are identified from those methods. The next step is using factor analysis to group those variables into 5 factors. In order to see the importance rank from these factors we will use AHP. The user can choose which antivirus is best for them based on the factors analysis and AHP method. This research will help common people in choosing the right antivirus for them.

1.1 Research Objective

1. Identifying which factors that need to be considered in choosing antivirus software by using factor analysis
2. Ranking the identified factors using AHP

1.2 Research Benefits

1. Helping the common people in deciding which antivirus software for them.
2. As an additional information for the IT business industry.

2 Literature Review

2.1. Factor Analysis

Factor analysis including both principal component analysis and common factor analysis is a statistical approach that can be used to analyzed interrelationships among a large number of variables and to explain these

* Corresponding author: ai.nurhayati@sttbandung.ac.id

variables in terms of their common underlying dimensions (actors) [6]. The objective is to find a way of condensing the information contained in a number of original variables into a smaller set of variates (factors) with a minimal loss of information [6]. For example, it is possible that variations in six observed variables mainly reflect the variations in two unobserved variables. Factor analysis searches for such joint variations in response to unobserved latent variables. Factor analysis aims to find independent latent variables. Followers of factor analytic methods believe that the information gained about the interdependencies between observed variables can be used later to reduce the set of variables in a dataset. The factor analysis model is scale invariant. The loadings are not unique (only up to multiplication by an orthogonal matrix) [18].

2.2 Analytic Hierarchy Process

Saaty [16] describes the seven pillars of the AHP as follows:

1. Ratio scales, proportionality and normalised ratio scales.
2. Reciprocal paired comparisons.
3. The sensitivity of the principal right eigenvector.
4. Clustering and using pivots to extend the scale.
5. Synthesis to create a one-dimensional ratio scale for representing the overall outcome.
6. Rank preservation and reversal.
7. Integrating group judgements.

The AHP provides a means of decomposing the problem into a hierarchy of subproblems which can more easily be comprehended and subjectively evaluated [10]. The subjective evaluations are converted into numerical values and processed to rank each alternative on a numerical scale [10]. The methodology of the AHP can be explained in following steps [10] :

Step 1: The problem is decomposed into a hierarchy of goal, criteria, sub-criteria and alternatives.

Step 2: Data are collected from experts or decision-makers corresponding to the hierarchic structure, in the pairwise comparison of alternatives on a qualitative scale.

Step 3: The pairwise comparisons of various criteria generated at step 2 are organised into a square matrix.

Step 4: The principal eigenvalue and the corresponding normalised right eigenvector of the comparison matrix give the relative importance of the various criteria being compared.

Step 5: The consistency of the matrix of order n is evaluated.

Step 6: The rating of each alternative is multiplied by the weights of the sub-criteria and aggregated to get local ratings with respect to each criterion.

3 Research Methodology

First we spread questionnaires to consumer, then from those questionnaires we identified 16 variables that needs to be considered on selecting antivirus software.

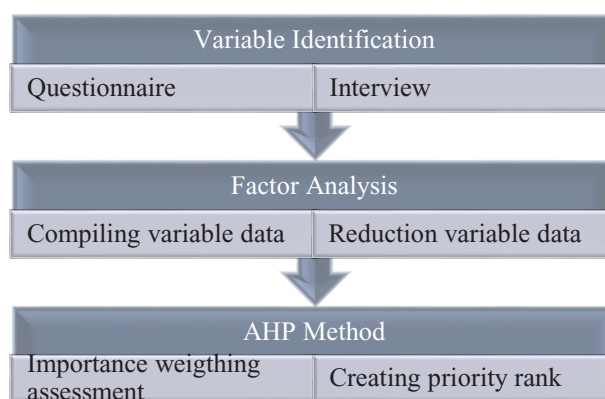


Fig 2. Flow chart of research

The 16 variables are anti-spyware, internet security, anti-trojan, anti-worm, anti-spam, online virus detection, software loading speed, user friendly, no excessive memory use, scanning virus fast and accurately, thorough scanning, affordable price, installing time, and data and identity protection. This variables then divided into factors by using factor analysis method in SPSS software. To rank those factors we spread questionnaires to 6 IT experts then the data is analyzed using AHP method. They all gained expertise in Anti Viruses for two years or certified.

4 Result and Review

First we spread questionnaires to consumer, then from those questionnaires we identified 16 variables that needs to be considered on selecting antivirus software.

4.1 Result from Factor Analysis

Output result from SPSS software program using Factor Analysis:

Table 1. KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.712
Bartlett's Test of Sphericity	Approx. Chi-Square	384.008
	Df	120
	Sig.	.000

Table 2. Rotated component Matrix

	Component				
	1	2	3	4	5
Software loading speed	.181	.744	.178	.020	.108
Software harddisk space	.118	.127	.032	.157	.839
Anti spyware	.673	.361	.193	.171	-.105
Internet security	.664	-.058	.188	.054	.236
Anti Trojan	.830	.300	-.008	.080	-.021
Anti worm	.879	-.011	-.103	-.057	.149
Installing time	.340	-.106	.077	.803	.162
Anti spam	.717	.063	-.062	.272	.323
User friendly	.188	.641	.138	-.205	.097
No excessive memory use	-.155	.871	.114	.206	-.045
Scanning virus fast and accurately	.241	.680	.024	-.282	.180
Thorough scanning	.179	.589	.180	-.487	.260
Virus scanning before download	.249	.345	.374	-.313	.607
Affordable Price	.110	.128	.819	.334	.126
Data and identity protection	.015	.240	.842	-.221	.047
Online virus detection	.614	.248	.362	-.179	-.036

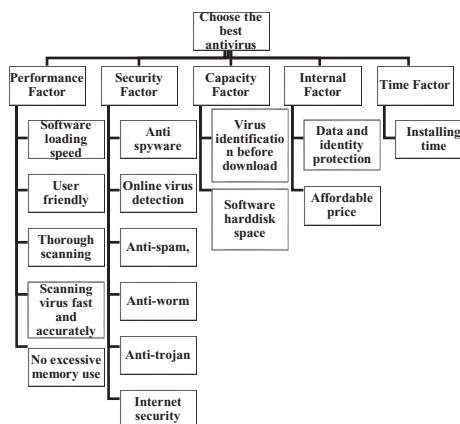


Fig 3. Hierarchy criteria

To rank those factors we spread questionnaires to 6 IT expert then the data is analyzed using AHP method.

Table 3. Expert 1

Expert 1	Security	Performance	Internal	Time	Capacity
Security	1	4	1/5	5	5
Performance	1/4	1	5	5	5
Internal	5	1/5	1	4	5
Time	1/5	1/5	1/4	1	3
Capacity	1/5	1/5	1/5	1/3	1

Table 4. Expert 2

Expert 2	Security	Performance	Internal	Time	Capacity
Security	1	1/9	5	5	5
Performance	9	1	9	9	9
Internal	1/5	1/9	1	1/5	1/5
Time	1/5	1/9	5	1	1/5
Capacity	1/5	1/9	5	5	1

Table 5. Expert 3

Expert 3	Security	Performance	Internal	Time	Capacity
Security	1	1/5	1/6	7	1/6
Performance	5	1	1/6	6	7
Internal	6	6	1	5	5
Time	1/7	1/6	1/5	1	1/5
Capacity	6	1/7	1/5	5	1

The 1st Factor variables are: Anti spyware, Internet security, Anti-trojan, Anti-worm, Anti-spam, Online virus detection. If we named this factor then it will be called Security Factor. This factor indicate that there are a group of consumer who bought antivirus software based on security factor. The 2nd Factor variables are: Software loading speed, User friendly, No excessive memory use, Scanning virus fast and accurately, Thorough scanning. If we named this factor then it will be called Performance Factor. The 3rd Factor variables are: Affordable price, Data and identity protection. This factor will be named Internal Factor. The 4th Factor variables are: Installing time. This factor will be named Time Factor. The 5th Factor variables are: Software harddisk space, Virus identification before download. This factor will be named Capacity Factor.

4.2 Result from AHP Calculations

The 16 variables then divided into 5 factors by using factor analysis. These five factors are security, performance, internal, time and capacity.

Table 6. Expert 4

Expert 4	Security	Performance	Internal	Time	Capacity
Security	1	1/3	1/3	1/3	1/4
Performance	3	1	1/4	1/4	1/3
Internal	3	4	1	1/3	1/3
Time	3	4	3	1	1/3
Capacity	4	3	3	3	1

Table 7. Expert 5

Expert 5	Security	Performance	Internal	Time	Capacity
Security	1	1/4	1/2	3	1/5
Performance	4	1	4	5	1/6
Internal	2	1/4	1	1/5	1/5
Time	1/3	1/5	5	1	1/6
Capacity	5	6	5	6	1

Table 8. Expert 6

Expert 6	Security	Performance	Internal	Time	Capacity
Security	1	7	1/5	1/3	1/8
Performance	1/7	1	8	8	1/6
Internal	5	1/8	1	9	7
Time	3	1/8	1/9	1	3
Capacity	8	1/6	1/7	1/3	1

Average assessment from 6 Experts:

$$GM_{12} = \sqrt[6]{4 \times \frac{1}{9} \times \frac{1}{5} \times \frac{1}{3} \times \frac{1}{4} \times 7} = 0,611$$

$$GM_{13} = \sqrt[6]{\frac{1}{5} \times 5 \times \frac{1}{6} \times \frac{1}{3} \times \frac{1}{2} \times \frac{1}{5}} = 0,421$$

$$GM_{14} = \sqrt[6]{5 \times 5 \times 7 \times \frac{1}{3} \times 3 \times \frac{1}{3}} = 1,969$$

$$GM_{15} = \sqrt[6]{5 \times 5 \times \frac{1}{6} \times \frac{1}{4} \times \frac{1}{5} \times \frac{1}{8}} = 0,544$$

$$GM_{21} = \sqrt[6]{\frac{1}{4} \times 9 \times 5 \times 3 \times 4 \times \frac{1}{7}} = 1,637$$

$$GM_{23} = \sqrt[6]{5 \times 9 \times \frac{1}{6} \times \frac{1}{4} \times 4 \times 8} = 1,979$$

$$GM_{24} = \sqrt[6]{5 \times 9 \times 6 \times \frac{1}{4} \times 5 \times 8} = 3,732$$

$$GM_{25} = \sqrt[6]{5 \times 9 \times 7 \times \frac{1}{3} \times \frac{1}{6} \times \frac{1}{6}} = 1,195$$

$$GM_{31} = \sqrt[6]{5 \times \frac{1}{5} \times 6 \times 3 \times 2 \times 5} = 2,376$$

$$GM_{32} = \sqrt[6]{\frac{1}{5} \times \frac{1}{9} \times 6 \times 4 \times \frac{1}{4} \times \frac{1}{8}} = 0,505$$

$$GM_{34} = \sqrt[6]{4 \times \frac{1}{5} \times 5 \times \frac{1}{3} \times \frac{1}{5} \times 9} = 1,157$$

$$GM_{35} = \sqrt[6]{5 \times \frac{1}{5} \times 5 \times \frac{1}{3} \times \frac{1}{5} \times 7} = 1,152$$

$$GM_{41} = \sqrt[6]{\frac{1}{5} \times \frac{1}{5} \times \frac{1}{7} \times 3 \times \frac{1}{3} \times 3} = 0,508$$

$$GM_{42} = \sqrt[6]{\frac{1}{5} \times \frac{1}{9} \times \frac{1}{6} \times 4 \times \frac{1}{5} \times \frac{1}{8}} = 0,268$$

$$GM_{43} = \sqrt[6]{\frac{1}{4} \times 5 \times \frac{1}{5} \times 3 \times 5 \times \frac{1}{9}} = 0,864$$

$$GM_{45} = \sqrt[6]{3 \times \frac{1}{5} \times \frac{1}{5} \times \frac{1}{3} \times \frac{1}{6} \times 3} = 0,521$$

$$GM_{51} = \sqrt[6]{\frac{1}{5} \times \frac{1}{5} \times 6 \times 4 \times 5 \times 8} = 1,837$$

$$GM_{52} = \sqrt[6]{\frac{1}{5} \times \frac{1}{9} \times \frac{1}{7} \times 3 \times 6 \times \frac{1}{6}} = 0,46$$

$$GM_{53} = \sqrt[6]{\frac{1}{5} \times 5 \times \frac{1}{5} \times 3 \times 5 \times \frac{1}{7}} = 0,868$$

$$GM_{54} = \sqrt[6]{\frac{1}{3} \times 5 \times 5 \times 3 \times 6 \times \frac{1}{3}} = 1,919$$

Table 9

	Security	Performance	Internal	Time	Capacity
Security	1	0,611	0,421	1,969	0,544
Performance	1,637	1	1,979	3,732	1,195
Internal	2,376	0,505	1	1,157	1,152
Time	0,508	0,268	0,864	1	0,521
Capacity	1,837	0,46	0,868	1,919	1
Total	7,358	2,844	5,132	9,777	4,412

Table 10

	Security	Performance	Internal	Time	Capacity	Total	Priority Weight
Security	0,136	0,215	0,082	0,201	0,123	0,757	0,151
Performance	0,222	0,352	0,386	0,382	0,271	1,613	0,323
Internal	0,323	0,178	0,195	0,118	0,261	1,075	0,215
Time	0,069	0,094	0,168	0,102	0,118	0,551	0,11
Capacity	0,25	0,162	0,169	0,196	0,227	1,004	0,201
Total	1	1	1	1	1	5	1

Calculation of matrix consistency test:

a. Calculate value [A]

$$[A] = \text{matrix} \times \text{priority weights} \quad (1)$$

$$A = \begin{bmatrix} 1 & 0,611 & 0,421 & 1,969 & 0,544 \\ 1,637 & 1 & 1,979 & 3,732 & 1,195 \\ 2,376 & 0,505 & 1 & 1,157 & 1,152 \\ 0,508 & 0,268 & 0,864 & 1 & 0,521 \\ 1,837 & 0,46 & 0,868 & 1,919 & 1 \end{bmatrix} \times \begin{bmatrix} 0,151 \\ 0,323 \\ 0,215 \\ 0,11 \\ 0,201 \end{bmatrix}$$

$$= \begin{bmatrix} 0,765 \\ 1,646 \\ 1,096 \\ 0,564 \\ 1,025 \end{bmatrix}$$

b. Calculate vector $B = \frac{\text{Vector [A]}}{\text{Priority weights}}$ (2)

$$B = \begin{bmatrix} 0,765 & 1,646 & 1,096 & 0,564 & 1,025 \\ 0,151 & 0,323 & 0,215 & 0,11 & 0,201 \end{bmatrix}$$

$$= |5,066 \quad 5,096 \quad 5,098 \quad 5,127 \quad 5,099|$$

c. Calculate Maximum Eigen Value

$$MEV = \frac{\text{Sum element on Matrix B}}{N} \quad (3)$$

$$MEV = \frac{5,066+5,096+5,098+5,127+5,099}{5} = 5,097$$

d. Calculate Consistency Index = $\frac{M.E.V-N}{N-1}$ (4)

$$\text{Consistency Index or CI} = \frac{5,097-5}{5-1} = 0,024$$

e. Search for random index value from table. If N = 5, then index random value = 1,12

f. Calculating and checking the Consistency Ratio.

$$\text{Consistency Ratio} = \frac{CI}{RI} \quad (5)$$

$$\text{Consistency Ratio} = \frac{0,024}{1,12} = CR = 0,021$$

From AHP calculations, we can conclude that the calculation result are consistent since CR = 0.021 (smaller than 0.1). AHP method also give us the importance rank from each of the factors. The most important factors are performance, followed by internal, capacity, security, and time.

The factors priority weight score are:

- Performance = 0,323
- Internal = 0,215
- Capacity = 0,201
- Security = 0,151
- Time = 0,11

4 Conclusions

The conclusion of the results of this research:

1. The 16 research variables are grouped to 5 factors. This is the result from Factor Analysis.
2. The factors and variables are :
 - a. Security factor are: Anti spyware, Internet security, Anti-trojan, Anti-worm, Anti-spam, Online virus detection.
 - b. Performance factor are: Software loading speed, User friendly, No excessive memory use, Scanning virus fast and accurately, Thorough scanning.
 - c. Internal factor are: Affordable price, Data and identity protection.
 - d. Time factor are: Installing time.

- e. Capacity factor are: Software harddisk space, Virus identification before download.
- 3. The factors priority weight score are:
 - a. Performance = 0,323
 - b. Internal = 0,215
 - c. Capacity = 0,201
 - d. Security = 0,151
 - e. Time = 0,11
- 4. Consumer can select antivirus software by judging the variables in the performance factors. Those variables are software loading speed, user friendly, no excessive memory use, thorough scanning, and scanning virus fast and accurately. Those variables can viewed at antivirus software.

References

1. D. K. Devi, Dr. K. M. Kumar, *An Analysis of Various Anti-Virus Software Tools Based On Different Effective Parameters*. ijctjournal, **4**(4) (2016)
2. F. Mamaghani, Evaluation and selection of an antivirus and content filtering software, *Information Management & Computer Security*, Vol. **10** Issue: 1, pp.28-32(2002)
3. F. Mata, J. L. Garcia-Dorado, J. Aracil, J. E. López de Vergara, Factor analysis of Internet traffic destinations from similar source networks, *Internet Research*, Vol. **22** Issue: 1, pp.29-56, (2012)
4. H. Chong, R. M. Zin, Selection of dispute resolution methods: factor analysis approach", *Engineering, Construction and Architectural Management*, Vol. **19** Issue: 4(2012)
5. I. Khan, An introduction to computer viruses: problems and solutions, *Library Hi Tech News*, Vol. **29** Issue: 7, pp.8-12(2012)
6. J. F. Hair et al. *Multivariate Data Analysis*. Pearson (2010)
7. M. Brunelli, *Introduction to the Analytic Hierarchy Process*. Springer International Publishing, Finland (2015)
8. M. Hosseini, M. S. Owlia, Designing a model for measuring and analyzing the relational capital using factor analysis: Case study, Ansar bank, *Journal of Intellectual Capital*, Vol. **17** Issue: 4, pp.734-757(2016)
9. M. Peters, S. Zelewski, Pitfalls in the application of analytic hierarchy process to performance measurement, *Management Decision*, Vol. **46** Issue: 7, pp.1039-1051(2008)
10. N. Bhushan, & K. Rai, *Strategic Decision Making Applying the Analytic Hierarchy Process*. Bangalore: Springer (2004)
11. R. Sriramachandramurthy, S.K. Balasubramanian, M. A. Hodis, Spyware and Adware: How Do Internet Users Defend Themselves?, *American Journal of Business*, Vol. **24** Issue: 2, pp.41-52(2009)
12. S. Chen, Y. Wu, L. Yang, Application of the Analytic Hierarchy Process for the selection of wastewater reuse targets, *Management Decision*, Vol. **52** Issue: 7, pp.1222-1235(2014)
13. S. Lin, M. Lu, Characterizing disagreement and inconsistency in experts judgments in the analytic hierarchy process, *Management Decision*, Vol. **50** Issue: 7, pp.1252-1265(2012)
14. S. Sipahi, M. Timor, The analytic hierarchy process and analytic network process: an overview of applications, *Management Decision*, Vol. **48** Issue: 5, pp.775-808(2010)
15. T. Saaty. *Decision making with the Analytic Hierarchy Process*. Int. J. Services Sciences Int. J. Services Sciences (2008)
16. T.L.Saaty, L.G.Vargas, *Models, Methods, Concepts and Applications of the Analytic Hierarchy Process*. Kluwer, Dordrecht (2001)
17. U. Roy, *Vulnerability of Watersheds to Climate Change Assessed by Neural Network and Analytical Hierarchy Process*. India: Springer (2016)
18. W. Hardley, L. Simar, *Applied Multivariate Statistical Analysis*. Springer (2003)
19. Wendler, T., & Gröttrup, S. *Factor Analysis*. Switzerland: Springer International Publishing (2016)