SCIENCE CHINA
Information Sciences

• Supplementary File •

A Revised CVSS-based System to Improve the Dispersion of Vulnerability Risk Scores

Chensi WU[1], Tao WEN[2] & Yuqing ZHANG[1,3*]

[1]*National Computer Network Intrusion Protection Center, University of Chinese Academy of Sciences, Beijing* 101408*, China;*
[2]*China Academy of Electronics and Information Technology, Beijing* 100041*, China;*
[3]*State Key Laboratory of Integrated Services Networks, Xidian University, Xi'an* 710071*, China*

Appendix A Paper details

The structure of this supplementary is: explaining the relevant work of security vulnerability risk assessment and CVSS; expounding the three criteria of metrics, and analyzes them based on CVSS; expounding the integral idea and implementation process of the CVSS_PCA; analyzing the dispersion and objectivity of CVSS_PCA.

Appendix A.1 Related work

CVSS 1.0 [1] was designed by NIST as a quantitative vulnerability risk assessment system in 2004. CVSS is acknowledged widely, and promotes the standardization process of the vulnerability evaluation [2]. In 2007, Mell et al. [3] upgraded the CVSS to 2.0 and made a comprehensive comparison between the two versions, they proved that the version 2.0 is better than 1.0 in both objectivity and dispersion. However, the objectivity and dispersion of 2.0 are still not perfect. CVSS 3.0 was put forward at the FIRST conference [4] in 2012, but it is still in testing phase so far.

In recent years, there is plentiful work [5] to improve the objectivity of CVSS. Ghani [6], Keramati [7], Fruhwirth [8], and Holm [9] researched CVSS metrics and put forward some attributes as new metrics, respectively. Wang [10] adjusted the formulas of CVSS. Temporal feature was added by Keramati [11]. Holm [12] manually marked 3000 vulnerabilities published on NVD to verify the objectivity of CVSS. Meanwhile, these vulnerabilities manually marked can be used as objective criteria of other quantitative systems. There is also some work about the dispersion of CVSS. Wang [13], Liu [14], Younis [15] et al. discussed the dispersion of CVSS in several. Based on CWE categories, Liu [16] et al. introduced category factor VTF and proposed VRSS 2.0 to improve the dispersion of CVSS 2.0 risk scores.

Appendix A.2 Analysis of CVSS dispersion and metrics

The dispersion of CVSS and the three criteria of metrics will be discussed below in detail.

Appendix A.2.1 *The distribution of CVSS*

There are some deficiencies in the dispersion of CVSS scores, for instance, six metrics of CVSS 2.0 have three possible values respectively,, that is there are total of $3^6 = 729$ results (729 risk scores). Unfortunately, most of the values are the same, and actually there are only 76 different scores, see Fig.A1. Meanwhile, the appearance probabilities of different risk scores are very uneven [3, 5]. Fig.A2 shows the dispersion of risk scores manually marked [12]. Obviously, it is a more even distribution and the pattern is closer to Normal Distribution.

Appendix A.2.2 *The probabilities of CVSS metric values*

40007 vulnerabilities published on NVD have been analyzed and Fig.A3 shows the probability of each metric value, for instance, the probabilities of L, N, A of metric Av is 9.68%, 85.69% and 4.63%, respectively. Clearly, the probabilities of each metric value are uneven, this lead to a poor dispersion.

Appendix A.2.3 *The total number of metric values*

Fig.A4 shows the distributions of CVSS metrics (VRSS metrics are the same), and Fig.A12 shows the distribution of CVSS_PCA. Obviously, the less number of metric values, the less number of risk scores.

[*] Corresponding author (email: zhangyq@nipc.org.cn)

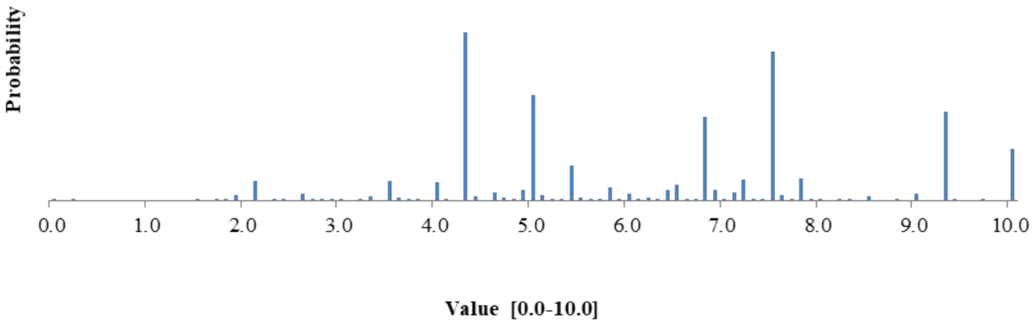


Figure A1 the distribution of CVSS 2.0 risk scores (40007 vulnerabilities).

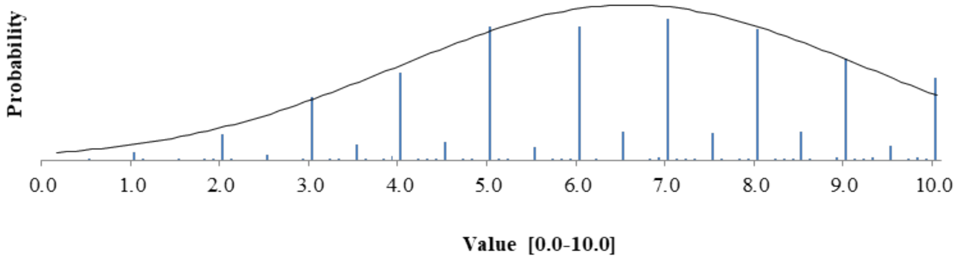
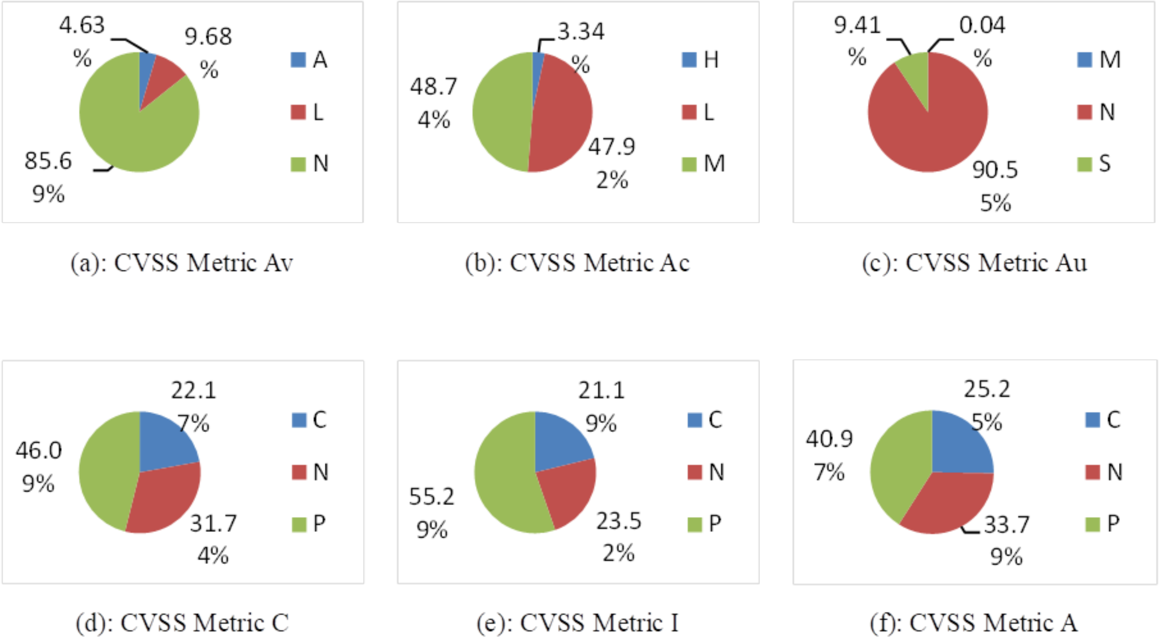


Figure A2 the distribution of risk scores manually marked (3000 vulnerabilities).



(a): CVSS Metric Av

(b): CVSS Metric Ac

(c): CVSS Metric Au

(d): CVSS Metric C

(e): CVSS Metric I

(f): CVSS Metric A

Figure A3 the distribution of risk scores manually marked (3000 vulnerabilities).

Appendix A.2.4 *The correlations between CVSS metrics*

The greater the difference between the conditional probability and the average probability becomes, the more obvious dependency is. Av is showed in detail as an example, and another five metrics are the same. Fig. A6 reveals the probabilities of three values of Av, in the case of each value of the other five metrics (five metrics * three values of each

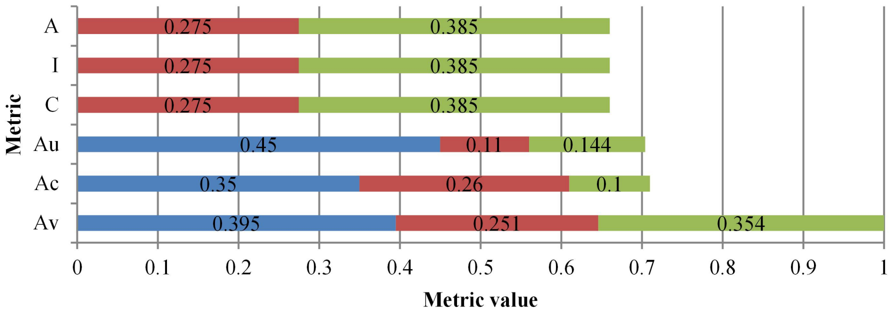
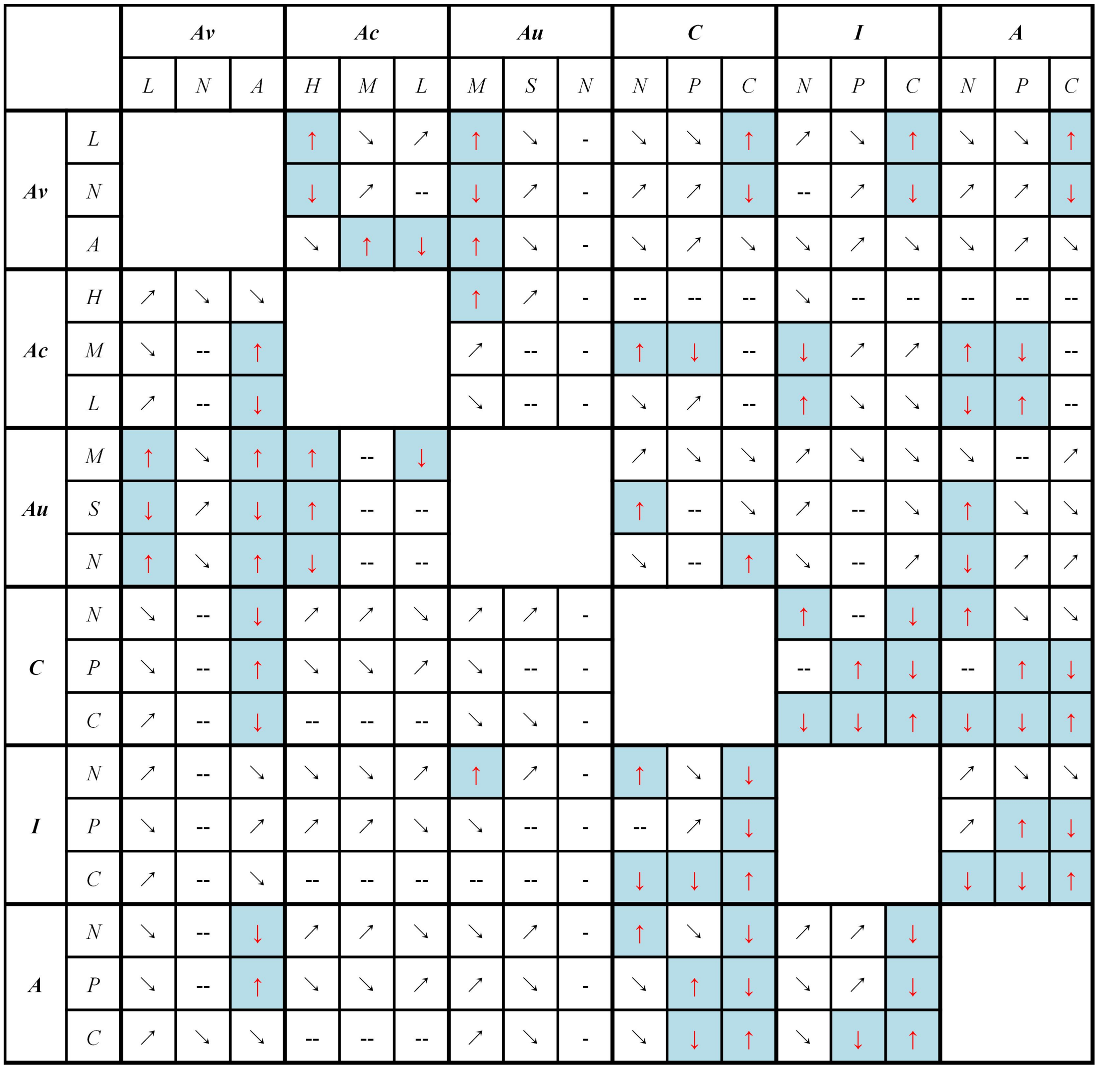


Figure A4 the distributions of CVSS and VRSS metrics.

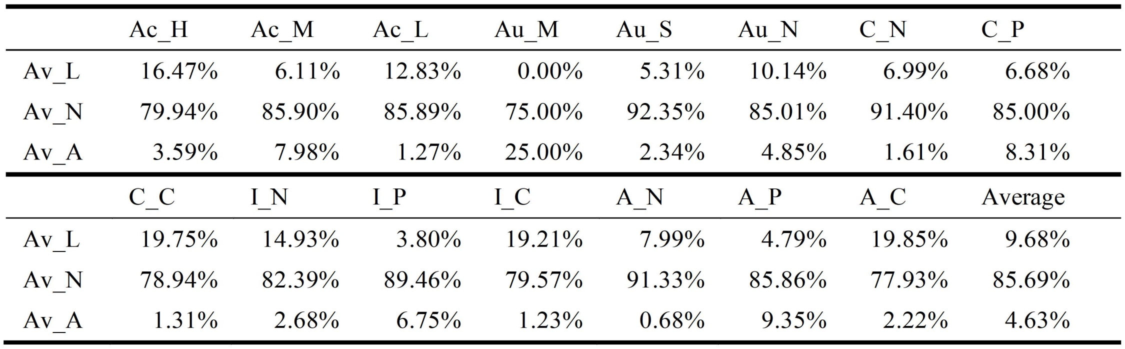


		Av			*Ac*			*Au*			*C*			*I*			*A*		
		L	*N*	*A*	*H*	*M*	*L*	*M*	*S*	*N*	*N*	*P*	*C*	*N*	*P*	*C*	*N*	*P*	*C*
Av	*L*				↑	↘	↗	↑	↘	-	↘	↘	↑	↗	↘	↑	↘	↘	↑
	N				↓	↗	--	↓	↗	-	↗	↗	↓	--	↗	↓	↗	↗	↓
	A				↘	↑	↓	↑	↘	-	↘	↗	↘	↘	↗	↘	↘	↗	↘
Ac	*H*	↗	↘	↘				↑	↗	-	--	--	--	↘	--	--	--	--	--
	M	↘	--	↑				↗	--	-	↑	↓	--	↓	↗	↗	↑	↓	--
	L	↗	--	↓				↘	--	-	↘	↗	--	↑	↘	↘	↓	↑	--
Au	*M*	↑	↘	↑	↑	--	↓				↗	↘	↘	↗	↘	↘	↘	--	↗
	S	↓	↗	↓	↑	--	--				↑	--	↘	↗	--	↘	↑	↘	↘
	N	↑	↘	↑	↓	--	--				↘	--	↑	↘	--	↗	↓	↗	↗
C	*N*	↘	--	↓	↗	↗	↘	↗	↗	-				↑	--	↓	↑	↘	↘
	P	↘	--	↑	↘	↘	↗	↘	--	-				--	↑	↓	--	↑	↓
	C	↗	--	↓	--	--	--	↘	↘	-				↓	↓	↑	↓	↓	↑
I	*N*	↗	--	↘	↘	↘	↗	↑	↗	-	↑	↘	↓				↗	↘	↘
	P	↘	--	↗	↗	↗	↘	↘	--	-	--	↗	↓				↗	↑	↓
	C	↗	--	↘	--	--	--	--	--	-	↓	↓	↑				↓	↓	↑
A	*N*	↘	--	↓	↗	↗	↘	↘	↗	-	↑	↘	↓	↗	↗	↓			
	P	↘	--	↑	↘	↘	↗	↗	↘	-	↘	↑	↓	↘	↗	↓			
	C	↗	↘	↘	--	--	--	↗	↘	-	↘	↓	↑	↘	↓	↑			

Figure A5 the correlations between CVSS metrics.

metric = 15), respectively. For instance, the probability is 16.47%, when Ac value is H and Av value is L. The average of Av_L is 9.68%, 16.47% is far larger than 9.68%, so it is a strong dependence between Ac_H and Av_L.

All the other metrics and qualitative correlations of values are shown in Fig. A5. Among them “↗” means strengthening. “↘” means weakening. “↑” means significant strengthening. “↓” means significant weakening. “–” means uncorre-lated. For example, when Ac value is H, the probability of Au value for M is significantly enhanced. From the table, A, C and I influence each other seriously, and because of the difference of conditional probabilities, the matrix is asymmetric.



	Ac_H	Ac_M	Ac_L	Au_M	Au_S	Au_N	C_N	C_P
Av_L	16.47%	6.11%	12.83%	0.00%	5.31%	10.14%	6.99%	6.68%
Av_N	79.94%	85.90%	85.89%	75.00%	92.35%	85.01%	91.40%	85.00%
Av_A	3.59%	7.98%	1.27%	25.00%	2.34%	4.85%	1.61%	8.31%
	C_C	I_N	I_P	I_C	A_N	A_P	A_C	Average
Av_L	19.75%	14.93%	3.80%	19.21%	7.99%	4.79%	19.85%	9.68%
Av_N	78.94%	82.39%	89.46%	79.57%	91.33%	85.86%	77.93%	85.69%
Av_A	1.31%	2.68%	6.75%	1.23%	0.68%	9.35%	2.22%	4.63%

Figure A6 the probability distribution of Av.

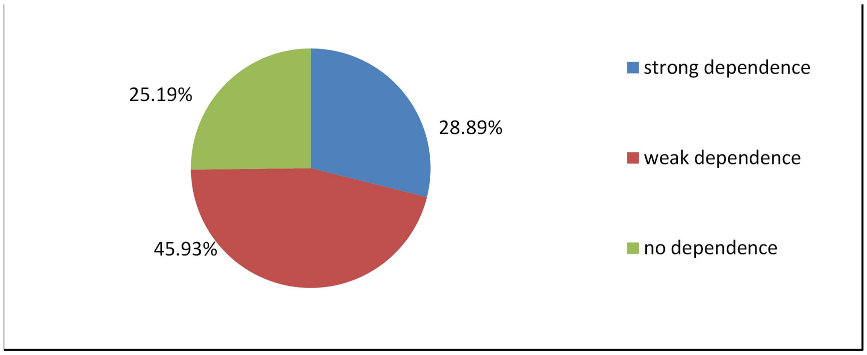


Figure A7 the statistics of CVSS metric correlations.

Fig.A7 shows the statistics of dependence between each metric, strong dependence (“↑” and “↓”) accounted for 28.89%, weak dependence (“↗” and “↘”) accounted for 45.93%, and no dependence (“–”) accounted for 25.19%. The correlations between the metrics may affect the probability distribution of metrics, which may further affect the risk score dispersion.

Appendix A.3 Introduction of CVSS_PCA

Appendix A.3.1 *Principal component analysis*

Principal Component Analysis is a statistical algorithm to compress and map metrics. PCA is often used in research of machine learning or data mining. It can make variables become dispersed, and try to keep the most original information [17]. The core of PCA is a linear combination of the original variables. In quantitative situation, it simplifies the original characteristics through linear transformation. Its prominent advantage is that it features on good decorrelation. PCA can put multiple mutual related metrics into several relatively independent metrics. In this paper, these characteristics of PCA are mainly used. The transformed principal component metrics are linear combinations of the original metrics.

PCA is introduced for improving CVSS. Assume that X is a variable with six dimensions{*dimensions: Av, Ac, Au, C, I, A*}, and $\mu = E(X)$, $\sum = Var(X)$.

$$\begin{cases} \mathrm{Z}_1 = a_1^T X \\ \mathrm{Z}_2 = a_2^T X \\ \mathrm{Z}_3 = a_3^T X \\ \mathrm{Z}_4 = a_4^T X \\ \mathrm{Z}_5 = a_5^T X \\ \mathrm{Z}_6 = a_6^T X, \end{cases}$$

Obviously,

$$\begin{aligned} Var(Z_i) &= a_i^T \sum a_i,\ i = 1, 2, ..., p, \\ Cov(Z_i, Z_j) &= a_i^T \sum a_j,\ i, j = 1, 2, ..., p, i \neq j. \end{aligned} \tag{A1}$$

The maximum variances of Z_1 is hoped to obtain, namely a_1 is the solution of constrained optimization problem,

$$max\ \ a^T \sum a\ and\ s.t.\ a^T a = 1.$$

Table A1 normalization range

Variable	Obtained range	Required range	Variable	Obtained range	Required range
CompAv	[0.39, 1.17]	[0.00, 1.00]	CompA	[-0.59, 0.51]	[0.00, 1.00]
CompAc	[-0.24, 0.56]	[0.00, 1.00]			
CompAu	[0.42, 0.91]	[0.00, 1.00]	Impact	[5.1, 10.4]	[0.0, 10.0]
CompC	[-0.14, 1.07]	[0.00, 1.00]	Exploitability	[0.0, 9.98]	[0.0, 10.0]
CompI	[-0.40, 0.57]	[0.00, 1.00]	Risk score	[0.8, 11.4]	[0.0, 10.0]

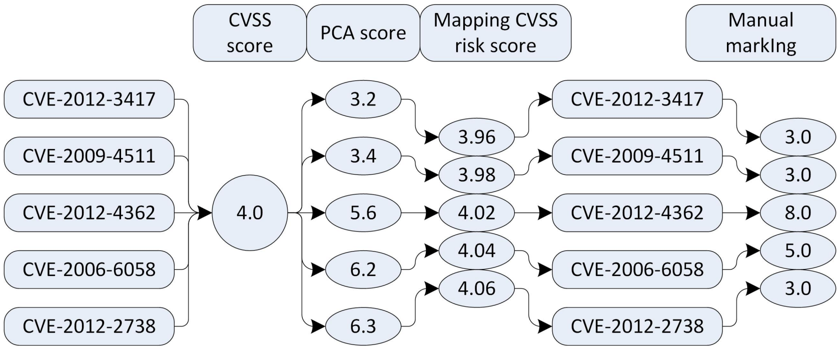


Figure A8 mapping CVSS risk scores with the order of PCA.

Therefore, a_1 is the feature vector (λ_1) of the largest eigenvalue of $\sum$. At this point, called $Z_1 = a_1^T \sum a_2 = 0$. Since a_1 is the eigenvector of λ_1, a_2 should be orthogonal to a_1. Similar to above deduction, a_1 is the feature vector (λ_2) of the second largest eigenvalue of $\sum$. $Z_2 = a_2^T X$ is called the second principal components. For covariance matrix $\sum$, normally, there is an orthogonal matrix Q. If we transform Q into a diagonal matrix, column i of Q will correspond to a_i and Z_i. In the implementation process of CVSS_PCA, PCA is used to calculate new metrics, which are {CompAv, CompAc, CompAu, CompC, CompI, CompA}. Each new metric is a linear combination of the initial metrics of CVSS. Formulas are shown in (A2)-(A7).

$$Comp_{Av} = Z_4 = 0.774Av + 0.597Ac - 0.148Au + 0.111A, \tag{A2}$$

$$Comp_{Ac} = Z_2 = -0.53Av + 0.763Ac + 0.32Au - 0.185I, \tag{A3}$$

$$Comp_{Au} = Z_3 = 0.313Av - 0.169Ac + 0.934Au, \tag{A4}$$

$$Comp_{C} = Z_1 = -0.144Av + 0.583C + 0.576I + 0.552A, \tag{A5}$$

$$Comp_{I} = Z_6 = 0.15Ac - 0.69C + 0.708I, \tag{A6}$$

$$Comp_{A} = Z_5 = -0.107Ac - 0.416C - 0.363I + 0.826A. \tag{A7}$$

Appendix A.3.2 *Normalization of variables*

For all the 40007 vulnerabilities, the obtained values of each variable are in a range, for example, CompAv is in [0.39, 1.17]. However, when a variable is used to input a function, the value range of this variable should be in the required range, for instance, the required range of CompAv is in [0.00, 1.00]. For other variables, see Table A1.

Appendix A.3.3 *Mapping of risk scores*

According to the order of PCA risk scores obtained in step 5, CVSS risk scores are mapped, and the vulnerabilities which are marked by same CVSS risk score can be marked by different risk scores. See in Fig.A8, the CVSS risk scores of the five vulnerabilities are 4.0, and the risk scores obtained by PCA are 3.2, 3.4, 5.6, 6.2, 6.3, respectively. Then we map the CVSS risk scores according to the order of PCA risk scores, and score interval is 0.02, so the results are 3.96, 3.98, 4.02, 4.04, 4.06. Note that the interval should be adjusted according to actual conditions.

Appendix A.4 Dispersion analysis based on experiment

In this section, CVSS_PCA is implemented and compared with CVSS 2.0 and VRSS 2.0 [16]. We compare not only the dispersion among three models, but also “the probability of metric values” and “the number of metric values”. It should be pointed out that “the correlations between metrics” do not be expounded, because the metrics transformed by PCA are less than all the other metric combinations in terms of correlation [18]. Note that:

- (For dispersion) Obtain all the vulnerabilities published on NVD, 70033 totally. Select the vulnerabilities which have been classified by CWE and assessed by CVSS 2.0. In fact, CVSS_PCA can be applied to all the vulnerabilities assessed by CVSS, but VRSS 2.0 only can be applied to the vulnerabilities classified by CWE. VRSS 2.0 is expected to make comparisons. only classified vulnerabilities are used, 40007 vulnerabilities totally.

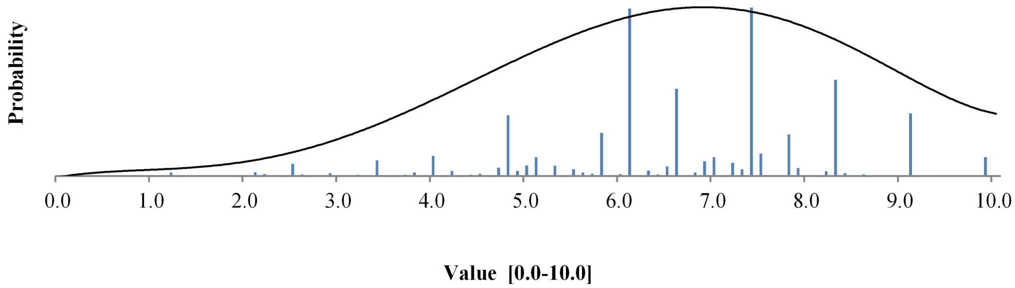


Figure A9 the distribution of risk scores of CVSS_PCA.

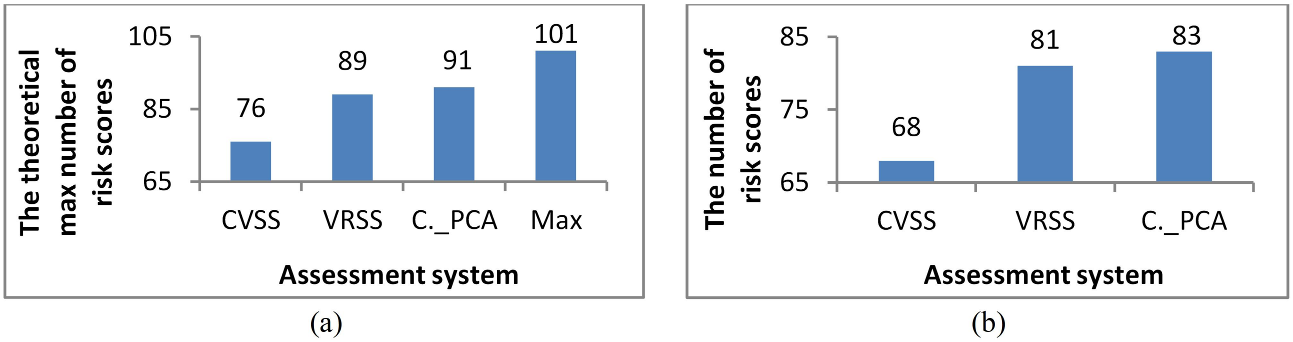


Figure A10 the number of risk scores.

- (For objectivity) Holm [2] assessed 3000 vulnerabilities published on NVD manually. Because the closer to the scores manually marked, the better objective, we regard these risk scores as objective assessment. Using VRSS to make a contrast, only 1882 classified vulnerabilities are marked manually.

Data shows that the dispersion of CVSS_PCA and VRSS 2.0 is better than CVSS 2.0. Considering application range, CVSS_PCA is better than VRSS 2.0.

Appendix A.4.1 *The distribution of CVSS_PCA*

Fig.A9 shows the distribution of risk scores of CVSS_PCA. Broadly, the distribution of CVSS_PCA risk scores is closer to the distribution of manual mark (see Fig.A2) than CVSS (see Fig.A1). The range of risk score is [0.0, 10.0], logically speaking, there are 101 risk values total. However, see Fig.A10 (a), the theoretical max number of risk scores of CVSS 2.0 is 76, VRSS 2.0 is 89, and CVSS_PCA is 91. In practice, within the set of 40007 vulnerabilities, the number of risk scores of CVSS 2.0 is 68, VRSS 2.0 is 81, and CVSS_PCA is 83, see Fig.A10 (b). With the increasing number of risk scores, the average number of vulnerabilities marked by same risk score is decreasing. So the dispersion of CVSS_PCA is better than CVSS. Although the difference of dispersion between CVSS_PCA and VRSS is not very apparent, CVSS_PCA on the objectivity and the performance is more excellent than that of VRSS. Meanwhile, the method of calculation of CVSS_PCA is simple and CVSS_PCA easies to do subsequent analysis using machine learning. VRSS needs to know the CWE category, which limits the application of VRSS.

Appendix A.4.2 *The probabilities of each metric value*

Fig.A11 shows the probabilities of each value of CompI, a CVSS_PCA metric. It can be seen from Fig.A3 that CompI has many values, and the difference of values between each other is smaller than CVSS metric I. The rest five CVSS_PCA metrics are in the similar situation.

Appendix A.4.3 *The total number of metric values*

Since the range of each metric value is [0.000, 1.000], the number of metric values can reach 1001. As shown in Fig.A12, actually, CVSS and VRSS are 12 (see Fig.A4. 0, 0.275, 0.35, 0.395, 0.45, 0.56, 0.61, 0.646, 0.66, 0.704, 0.71, and 1), CVSS_PCA is 301. Fig.A13 shows the distribution of CVSS_PCA metric values. The number of metric values of CVSS_PCA is more than CVSS and VRSS.

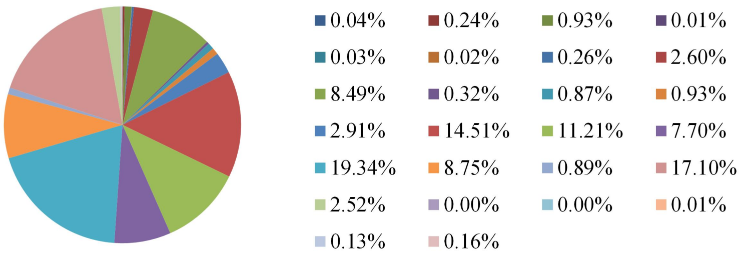


Figure A11 the probabilities of each value of CVSS_PCA metric CompI.

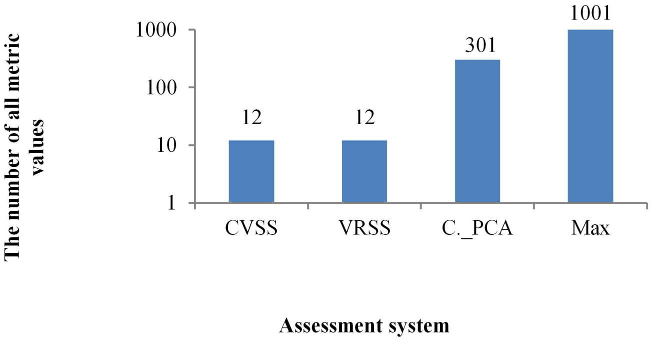


Figure A12 the number of all metric values.

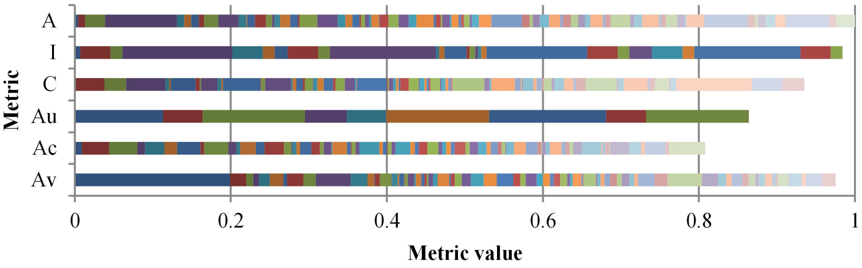


Figure A13 the distribution of CVSS_PCA metric values.

Appendix A.5 Objectivity analysis based on experiment

Based on 3000 vulnerabilities marked manually, which we regard as objective assessment, we use 1882 vulnerabilities of manual mark of VRSS as contrast data. Fig.A14 shows the distribution of the difference with risk scores manually marked. When the differences from risk scores manually marked are less than 2, there are 1383 vulnerabilities in CVSS_PCA, 1339 in CVSS, and 1112 in VRSS. Consequently, there is no obvious difference on objectivity among CVSS, VRSS, and CVSS_PCA. In order to further verify the objectivity, the variances of the difference in manual risk scores are calculated, see Fig.A15. The X axis is the interval or the amplitude of translation. The definition of interval is as shown in Fig.A8. When the interval is 0.9, the variance of the difference between CVSS_PCA and manual noted scores reaches the minimum. In order to verify that the decline of variance is mainly contributed by the change of interval, instead of translating, we translate CVSS risk scores. See Fig.A15, when the translation is 0.7, namely each CVSS risk score minus 0.7, the variance of the difference is minimized. But it is still bigger than the variance between CVSS_PCA risk scores and manual scores. For verifying the importance of mapping order, CVSS scores are mapped with a reversed PCA order, see Fig.A16 contrasting Fig.A8. The variance of the difference is increasing with the increase of interval, and all the variances are bigger than CVSS_PCA. Based on the above discussion, we can draw the conclusion that CVSS_PCA risk scores (mapping CVSS according to the order of PCA risk scores) are most close to scores manually marked.

References

1 Mell P, Scarfone K, Romanosky S. Common Vulnerability Scoring System. IEEE Security & Privacy, 2007, 4(6):85-89

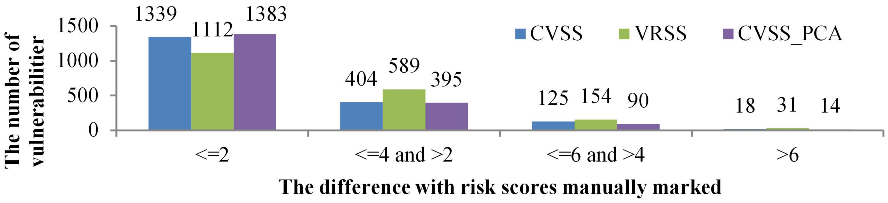


Figure A14 the distribution of the differences with manual risk scores.

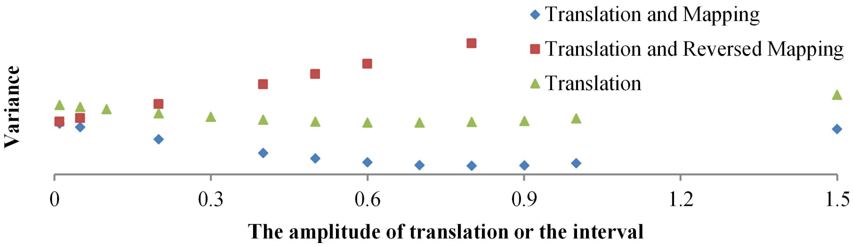


Figure A15 the variances of the difference with manual risk scores.

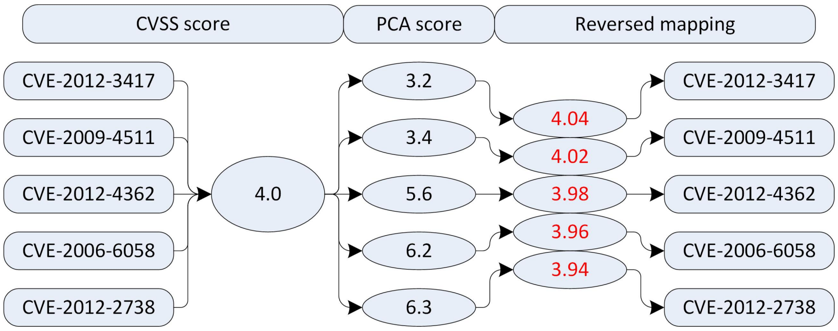


Figure A16 mapping CVSS risk scores with the reversed order of PCA.

2 Alhazmi O H, Malaiya Y K. Quantitative Vulnerability Assessment Of Systems Software. In: Reliability and Maintainability Symposium, 2005. 615-620

3 Scarfone K, Mell P. An Analysis Of CVSS Version 2 Vulnerability Scoring. In: Interna-tional Symposium on Empirical Software Engineering and Measurement (ESEM), Lake Buena Vista, 2009. 516-525

4 First Improving Security Together, https://www.first.org/cvss

5 Mell P, Scarfone K. Improving the Common Vulnerability Scoring System. Information Security Iet, 2007, 1(3):119-127

6 Ghani H, Luna J, Suri N. Quantitative Assessment Of Software Vulnerabilities Based On Economic-Driven Security Metrics. In: Risks and Security of Internet and Systems (CRi-SIS), International Conference, La Rochelle, 2013. 1-8

7 Keramati M. Novel Security Metrics For Ranking Vulnerabilities In Computer Networks. In: Telecommunications (IST), 2014 7th International Symposium on. Tehran, 2014. 883-888

8 Fruhwirth C, Mannisto T. Improving CVSS-Based Vulnerability Prioritization And Response With Context Information. In: Empirical Software Engineering and Measurement (ESEM), International Symposium, Lake Buena Vista, 2009. 535-544

9 Holm H, Ekstedt M, Andersson D. Empirical Analysis of System-Level Vulnerability Metrics through Actual Attacks. IEEE Transactions on Dependable & Secure Computing, 2012, 9(6):825-837

10 Wang L.Y, Jajodia S, Singhal A, et al. K-Zero Day Safety A Network Security Metric For Measuring The Risk Of Unknown Vulnerabilities. Dependable and Secure Computing, IEEE Transactions, 2013, 11:30-44

11 Keramati M.: New Vulnerability Scoring System for dynamic security evaluation. In: 2016 8th International Symposium on Telecommunications, 2017. 746-751

12 Holm H, Afridi K K. An expert-based investigation of the Common Vulnerability Scoring System. Computers and Security, 2015, 53:18-30

13 Wang R Y, Gao L, Sun D H. An Improved CVSS-Based Vulnerability Scoring Mechanism. In: Multimedia Information Networking and Security (MINES), International Conference, Shanghai, 2011. 352-355

14 Liu Q X, Zhang Y Q. VRSS: A New System For Rating And Scoring Vulnerabilities. Computer Communications, 2011, 34:264-273

15 Younis A A, Malaiya Y K, Ray I. Using Attack Surface Entry Points And Reachability Analysis To Assess The Risk Of Software vulnerability Exploitability. In: High-Assurance Systems Engineering (HASE), 2014 IEEE 15th International Symposium on Miami Beach, 2014. 1-8

16 Liu Q X, Zhang Y Q, Kong Y. Improving VRSS-Based Vulnerability Prioritization Using Analytic Hierarchy Process. The Journal of Systems and Software, 2012, 85:1699-1078

17 Hadri A, Chougdali K, Touahni R. Intrusion Detection System using PCA and Fuzzy PCA Techniques. In: 2016 International Conference on Advanced Communication Systems and Information Security, 2016. 1-7

18 Zhao L H, Guo Z K. Face Recognition Method Based on Adaptively Weighted Block-Two Dimensional Principal Component Analysis. In: Computational Intelligence, Communication Systems and Networks (CICSyN), 2011 Third International Conference on Bali, 2011. 22-25