Computer Science and Application 计算机科学与应用, 2012, 2, 121-125
http://dx.doi.org/10.12677/csa.2012.23022 Published Online August 2012 (http://www.hanspub.org/journal/csa.html)

A Metric Model Research Based on Attributes for Trustworthiness of Software*

Yanzhao Liu[1], Xun Luo[2], Kai Xue[3], Ping Luo[3]

[1]China Information Technology Security Evaluation Center, Beijing
[2]The Department of Electronic and Computer Engineering, The Hong Kong University of Science and Technology, Hong Kong
[3]Key Laboratory for Information System Security, Ministry of Education, Tsinghua National Laboratory for Information Science and Technology, School of Software, Tsinghua University, Beijing
Email: luop@mail.tsinghua.edu.cn

Received: Apr. 22nd, 2012; revised: May 11th, 2012; accepted: May 17th, 2012

Abstract: In recent years, trustworthiness of the software has become a frontier research in the field of software engineering. In this paper, the impact factor which is used to control the influence of the critical attribute with the minimal value in the metric model based on attributes is improved, and a new metric model, the impact factor being the function of the minimal value and its weight in the metric model, is proposed. This new model is more consistent with the practice and has its advantages.

Keywords: Trustworthiness of Software; Metric Model

一种基于属性划分的软件可信性度量模型研究*

刘彦钊[1]，罗　峋[2]，薛　凯[3]，罗　平[3]

[1]中国信息安全测评中心，北京
[2]香港科技大学，电子与计算机工程系，香港
[3]清华大学软件学院，信息系统安全教育部重点实验室，清华信息科学与技术国家实验室，北京
Email: luop@mail.tsinghua.edu.cn

收稿日期：2012 年 4 月 22 日；修回日期：2012 年 5 月 11 日；录用日期：2012 年 5 月 17 日

摘　要：近年来软件的可信性已经成为软件工程领域的一个前沿研究方向，本文对基于属性划分的软件可信性度量模型中控制具有最小度量值的关键子属性的影响因子进行了改进，提出了一种新的度量模型，使其影响因子是最小度量值和其权重的函数。该模型更加符合实际情况且具有其优越性。

关键词：软件可信性；度量模型

1. 引言

随着计算机技术和互联网技术的发展，计算机软件已经广泛应用于各个行业领域，各类应用软件系统已经渗透到人类社会的政治、经济、文化以及人们生活的各个层面，使得人们对软件的依赖程度越来越高。由于现代的软件系统充分考虑了软件最终用户的各种需求，交互性很强，这往往导致它是一个复杂巨系统，从而与之产生的问题就是软件缺陷和软件漏洞呈爆炸式增长，并由此带来的重大事故和严重损失屡见不鲜：1962 年 7 月 22 日，带有飞越金星任务的水手 1 号火箭在发射升空 5 分钟左右的时候偏离轨道，为了不造成更大的损失，美国空军将其摧毁，之后的调查小组发现，事故的起因仅仅是因为某程序员将草稿上的公式抄写成了错误的程序代码。还有 2003 年造成直接损失 60 亿美元的北美大停电事件，也是由于美国电力控制系统的软件失效造成的。

*资助信息：本文得到国家自然科学基金重点项目资助，项目号：90818021。

另外，随着网络的广泛使用，互联网软件的用户数量急剧增长，黑客软件日益强大，而网络的开放性使得黑客工具和教程更加容易获取，这对网络软件的可靠性和安全性等属性提出了更高的要求，而这些问题本质上是软件的可信性问题，即什么样的软件才会被人们所信任，可以称之为可信软件，信任的程度又如何进行度量等等，本文认为这两个问题正是软件可信性理论的核心问题。它迫使许多国内外学者们研究软件应用中出现的可信性问题，纷纷要求对软件系统的可信性提出度量。

本文第 2 节将介绍软件可信性度量的研究背景以及软件可信性度量函数的四条性质和基于属性划分的一种度量模型[1]。第 3 节将深入分析该模型，并基于此提出了一种改进的度量模型。第 4 节针对本文提出的改进模型给出了一个简单算例，说明改进模型的正确性和优越性。第 5 节总结全文并展望下一步的工作。

2. 软件可信性的度量

上世纪 70 年代，Anderson 首次提出了可信系统的概念[2]，自此学术界与工业界都从不同的角度对信息系统的可信性做出了各自的表述。ISO/IEC 15408 将系统可信性[3]定义为：一个可信的组件、操作或过程的行为，在任意操作条件下是可以预测的，并能很好地抵抗应用软件、病毒以及一定的物理干扰造成的破坏。可信计算组织则认为，一个实体如果总是按照其设定的目标所期望的方式运行，则这个实体是可信的。Algirdas[4]则把传统软件的可信性分为可靠性与安全性两个方面。王怀民[5]等人则将软件系统可信性总结为身份可信与能力可信。

由上可知，虽然目前学术界对软件可信性问题众说纷纭，并没有一个统一确定的标准定义，其核心意义含糊、笼统，不够深刻。但一般认为软件可信性是建立在一些子属性(可靠性、安全性、可用性等)的基础上的一个新概念，可信子属性的划分涉及了软件工程需求分析，并不是本文的讨论重点，本文的目的在于分析基于子属性划分的软件可信性应该如何度量，即假设软件可信性是建立在一个可信属性集合上的，我们研究其度量问题。事实上如果搞清楚了这个问题，那么对于已经有成熟的、有明确需求要求的行业软件来说，我们就可以度量其软件系统是否可信，并通过定量计算，比较不同软件之间的可信度差异，这也正是本文的应用意义所在。

文[1]中将软件的可信度量值定义为各个可信子属性度量值的函数：

$$T = T\left(y_1, y_2, \cdots, y_n\right) \tag{1}$$

其中 y_i 表示软件可信子属性的第 i 个度量值变元，T 表示软件的可信度，并且有 $y_i \in [1,10]$，$T \in [1,10]$。对于一个软件系统，可将其可信子属性分为两类：关键属性与非关键属性。关键属性指的是一个可信软件所必须具有的基本属性，比如可靠性、安全性、正确性等。如果软件系统的任意一个关键属性的度量值低于阈值，那么认为该软件不可信。由于各个软件的应用场合与用户的不同，很多软件具有另外的一些属性，例如可维护性，可移植性等，这类属性则称之为非关键属性。关键属性与非关键属性是相对的，例如对于一个数值计算软件来说，正确性是关键属性，而界面友好性是非关键属性；但对于一个面向大众网民的网络软件系统，界面友好性则是关键属性，正确性反而下降为非关键属性。

为了讨论方便起见，设 $y_1, y_2, \cdots, y_m$ 表示为关键属性，$y_{m+1}, y_{m+2}, \cdots, y_{m+s}$ 为非关键属性$(m+s=n)$。文献[1]提出了可信度量函数应该满足单调性，增长性，敏感性和替代性四条性质，并基于此提出了一个度量模型如下：设 $u_1 = \min\limits_{1 \le i \le m}\left\{\left(\frac{y_i}{10}\right)^{\varepsilon}\right\} y_1^{\alpha\alpha_1} y_2^{\alpha\alpha_2} \cdots y_m^{\alpha\alpha_m}$ 和 $u_2 = y_{m+1}^{\beta\beta_{m+1}} y_{m+2}^{\beta\beta_{m+2}} \cdots y_{m+s}^{\beta\beta_{m+s}}$，

$$T = \frac{10}{11}\left(u_1 + u_2\right) \tag{2}$$

其中 α 与 β 分别表示关键属性与非关键属性的权重，有 $\alpha + \beta = 1$ 且 $\alpha > 0.5 > \beta$；α_i 表示第 i 个关键属性在整个关键属性集中所占权重，有 $\sum\limits_{i=1}^{m} \alpha_i = 1$，对应的 β_i 表示第 i 个非关键属性在整个非关键属性集中所占权重，有 $\sum\limits_{i=m+1}^{m+s} \beta_i = 1$。$\varepsilon$ 满足 $0 \le \varepsilon \le 1$ 且 $\alpha\alpha_{\min} + \varepsilon < 1$，其中 $\alpha_{\min}$ 表示关键属性中具有最小度量值的权值。

3. 一种改进的可信性度量模型

上节简单介绍了一种基于属性划分的软件可信

性度量模型，本节在深入分析该模型的基础上，将提出一种改进的模型。

文献[1]证明了式(2)度量模型基本满足上节所介绍的四条性质：单调性，增长性，敏感性和替代性，比较符合实际情况。另外，由式(2)可以看出，该模型由于加入了 $\min\limits_{1\le i\le m}\left\{\left(\frac{y_i}{10}\right)^{\varepsilon}\right\}$ 一项，突出了具有最小度量值的可信关键属性对整个软件可信度的影响，而这个影响程度是由 ε 这一参数控制的，即对同一组子属性度量值不变的情况下，ε 越大，软件可信度越小，反之，软件的可信度越大。直观上来看，可以用“短板效应”来解释，如果把软件可信子属性比作 n 块木板，软件可信度当作是这 n 块木板围成的“木桶”的盛水量，那么该“木桶”的盛水量显然会被最短板所限制，而这个限制程度正由 ε 来控制。

原模型(2)中的 ε 是区间[0, 1]上的一个给定值，而这往往与实际情况不符。本文认为在现实应用中，对于不同的子属性度量值的取值，ε 应该是不同的，而不是固定不变且相等的。最小度量值越小，造成的“短板效应”越明显，对整个软件可信度的影响应该越大。另一方面，具有最小度量值的子属性的权重大小也应该直接影响可信度量值，也就是说，该权重越大，ε 应该越大，对可信度的影响越大，反之，该权重越小，ε 越小。由此可知原模型中的参数 ε 应该是一个关于 $y_{\min}$ 与 $\alpha_{\min}$ 的函数值，本文正是基于此提出了一种改进模型。

设 $u_3=\min\limits_{1\le i\le m}\left\{\left(\frac{y_i}{10}\right)^{k_1\alpha_i-k_2y_i}\right\}y_1^{\alpha\alpha_1}y_2^{\alpha\alpha_2}\cdots y_m^{\alpha\alpha_m}$，基于软件可信性度量模型(2)，我们给出如下度量模型函数：

$$T=\frac{10}{11}\left(u_3+u_2\right) \tag{3}$$

其中，$k_1\le 10/\alpha_{\min}\left(10-\alpha_{\min}\right)$，$k_2\le k_1\alpha_{\min}/10$ 且 $k_2\le\frac{(\alpha+k_1)\alpha_{\min}}{10(1+\ln 10)}$。$k_1$ 用来控制具有最小度量值的可信属性的权重对可信度的影响，k_2 用来控制最小度量值对可信度的影响。函数中其它参数意义同原模型保持一致。

定理：度量函数(3)满足单调性，增长性，敏感性和可替换性。

由于度量函数(3)加入了 $\left(\frac{y_i}{10}\right)^{k_1\alpha_i-k_2y_i}$ 项，因此可替换性的证明相对比较复杂。由于改进模型的重点在于强调“短板效应”，因此我们下面将给出可信度量函数(3)满足单调性，增长性，敏感性三条性质的证明，其第四条性质可替换性的证明可类似得到。

证明：a) T 是单调递增函数

记 $\min_{1\le i\le m}\{y_i\}$ 的下标 i 为 min，则 $y_{\min}=\min_{1\le i\le m}\{y_i\}$。又记

$$\chi_1=y_1^{\alpha\alpha_1}y_2^{\alpha\alpha_2}\cdots y_{\min}^{(\alpha+k_1)\alpha_{\min}-k_2y_{\min}}\cdots y_m^{\alpha\alpha_m},$$

$$\chi_2=\frac{(\alpha+k_1)\alpha_{\min}}{y_{\min}}-(1+\ln y_{\min})k_2+k_2\ln 10,$$

$$\chi_3=10^{k_1\alpha_{\min}-k_2y_{\min}},$$

$$\chi_4=y_1^{\alpha\alpha_1}y_2^{\alpha\alpha_2}\cdots y_i^{\alpha\alpha_i-1}\cdots y_{\min}^{(\alpha+k_1)\alpha_{\min}-k_2y_{\min}}\cdots y_m^{\alpha\alpha_m},$$

$$\chi_5=y_{m+1}^{\beta\beta_{m+1}}y_{m+2}^{\beta\beta_{m+2}}\cdots y_i^{\beta\beta_i-1}\cdots y_{m+s}^{\beta\beta_{m+s}},$$

对函数 T 关于 y_i 求偏导，得：

$$\frac{\partial T}{\partial y_i}=\begin{cases}\dfrac{\chi_1\times\chi_2}{1.1\times\chi_3}, & i=\min\\ \dfrac{\alpha\alpha_i\chi_4}{1.1\times\chi_3}, & i\ne\min,1\le i\le m\\ \dfrac{\beta\beta_i\chi_5}{1.1}, & m+1\le i\le m+s\end{cases}$$

由于 $k_2\le\frac{(\alpha+k_1)\alpha_{\min}}{10(1+\ln 10)}$，有 $\frac{(\alpha+k_1)\alpha_{\min}}{y_{\min}}-(1+\ln y_{\min})k_2\ge 0$，所以 $\frac{\partial T}{\partial y_i}\ge 0$，因此 T 是单调递增函数。

b) 函数 T 的值域是[1, 10]

由上面的证明可知：因为 T 是单调递增函数，又有 $1\le y_i\le 10\left(1\le i\le m+s\right)$，因此可得：

$$\frac{10}{11}\left(\frac{y_{\min}}{10}\right)^{k_1\alpha_{\min}-k_2y_{\min}}+\frac{10}{11}\le T\le\frac{10}{11}10^{\alpha}+\frac{10}{11}10^{\beta}$$

由 $k_1\le 10/\alpha_{\min}\left(10-\alpha_{\min}\right)$，$k_2\le k_1\alpha_{\min}/10$，可知 $0\le k_1y_{\min}-k_2\alpha_{\min}\le 1$，则：

$$T\ge\frac{10}{11}\left(\frac{1}{10}\right)^{k_1y_{\min}-k_2\alpha_{\min}}+\frac{10}{11}\ge 1$$

令 $f(\alpha)=\frac{10}{11}10^{\alpha}+\frac{10}{11}10^{1-\alpha}\ (0.5<\alpha<1)$，易知

$f(\alpha)$是关于α的单调增函数，因此有$f(\alpha)\le 10$。故推得$1\le T\le 10$。

c) 函数T满足增长性

设：

$$\chi_6=\frac{k_2}{y_{\min}}+\frac{(\alpha+k_1)\alpha_{\min}}{y_{\min}^2},$$

$$\chi_7=y_1^{\alpha\alpha_1}y_2^{\alpha\alpha_2}\cdots y_i^{\alpha\alpha_i-2}\cdots y_{\min}^{(\alpha+k_1)\alpha_{\min}-k_2y_{\min}}\cdots y_m^{\alpha\alpha_m},$$

$$\chi_8=y_{m+1}^{\beta\beta_{m+1}}y_{m+2}^{\beta\beta_{m+2}}\cdots y_i^{\beta\beta_i-2}\cdots y_{m+s}^{\beta\beta_{m+s}},$$

对T求关于y_i的二阶偏导数，则有

$$\frac{\partial^2T}{\partial y_i^{\,2}}=\begin{cases}\dfrac{\chi_1}{1.1\times\chi_3}\left(\chi_2^2-\chi_6\right), & i=\min\\ \dfrac{\alpha\alpha_i(\alpha\alpha_i-1)\chi_7}{1.1\chi_3}, & i\ne\min,1\le i\le m\\ \dfrac{\beta\beta_i(\beta\beta_i-1)\chi_8}{1.1}, & m+1\le i\le m+s\end{cases}$$

当$1\le i\le m+s$时，$\dfrac{\partial^2T}{\partial y_i^{\,2}}$是一个分段函数，且其符号取决于参数$k_1$与$k_2$的取值，由于

$k_1\le 10/\alpha_{\min}(10-\alpha_{\min})$，$k_2\le k_1\alpha_{\min}/10$且

$k_2\le\dfrac{(\alpha+k_1)\alpha_{\min}}{10(1+\ln 10)}$，因此有$\dfrac{\partial^2T}{\partial y_i^{\,2}}\le 0$。

d) 函数T满足敏感性

通过计算，可得可信度T对每个可信子属性的敏感性为：

$$\frac{\partial T}{\partial y_i}\frac{y_i}{T}=\begin{cases}\dfrac{\chi_1\times\chi_2\times y_i}{1.1\times\chi_3\times T}, & i=\min\\ \dfrac{\alpha\alpha_i\chi_4\times y_i}{1.1\times\chi_3T}, & i\ne\min,1\le i\le m\\ \dfrac{\beta\beta_i\chi_5y_i}{1.1T}, & m+1\le i\le m+s\end{cases}$$

从上式可以看出，由于实际中可信关键属性的权重α远大于非关键属性的权重β，因此T对关键属性的敏感度要高于非关键属性。此外，该式反应了可以通过调节参数k_1与k_2的大小来控制最小关键属性的度量值与权重对T的影响程度，k_1越大，最小关键属性的权重对T影响越大，k_2越小，最小关键属性的度量值对T的影响越大。

4. 简单的应用实例

为了验证我们上面提出的改进模型的正确性和优越性，本节对软件可信性度量模型(3)给出了一个简单实例，设$m=3$，$s=1$。取$\alpha=0.9$，$\beta=0.1$，根据不同的子属性度量值及其权重的取值，给出计算结果如表1所示。

从表1的七组数据对比中，可以得出如下结论：

1) T对任意的y_i都是单调递增的，但显然从前四组数据对比中可以看到，T对关键属性的敏感程度要大于非关键属性，而对具有最小度量值的关键属性(y_3)的敏感度则最高。

2) 对比第1组数据与第5组数据，当具有最小度量值的y_3的权重从0.3调整为0.2时，按照模型的原意，在参数k_2的控制下，可信度量值应该是低于4.78的，但由于在权重调整时，同时将关键属性y_2的权重调整为了0.5，因此得到5.55的结果，高于5.95，这是符合软件运行实际情况的。

3) 最后两组数据分别反应了参数k_1与k_2的作用，k_1越大，最小关键属性的权重对T的影响度越高，k_2越小，最小关键属性的度量值对T的影响越大，因此对照第1组数据，当k_1，k_2分别调整为3.4与0.01时，T度量值都有了不同程度的下降。

取$\varepsilon=0.78$，$k_1=3.2$，$k_2=0.03$，表2对原模型(2)和改进后的模型(3)做了比较。

表2中T_1与T_2表示分别应用原模型与改进后的模型计算得到的软件可信度量值，对于第1组数据，

Table 1. Experimental results between T and the parameters
表1. T与各参数关系实验结果

	y_1/α_1	y_2/α_2	y_3/α_3	y_4/β_1	k_1	k_2	T
1	8/0.3	8/0.4	6/0.3	8/1	3.2	0.03	4.78
2	8/0.3	9/0.4	6/0.3	8/1	3.2	0.03	4.94
3	8/0.3	8/0.4	6/0.3	9/1	3.2	0.03	4.80
4	8/0.3	8/0.4	7/0.3	8/1	3.2	0.03	5.00
5	8/0.3	8/0.5	6/0.2	8/1	3.2	0.03	5.55
6	8/0.3	8/0.4	6/0.3	8/1	3.4	0.03	4.67
7	8/0.3	8/0.4	6/0.3	8/1	3.2	0.01	4.57

Table 2. Results of the comparison of the original model T_1 and improved model T_2
表2. 原模型T_1与改进模型T_2比较结果

	y_1/α_1	y_2/α_2	y_3/α_3	y_4/β_1	T_1/T_2
1	8/0.3	8/0.4	6/0.3	8/1	4.78/4.78
2	8/0.3	8/0.4	5/0.3	8/1	4.61/4.56
3	8/0.3	8/0.4	6/0.35	8/1	5.56/5.2

易得 $k_1\alpha_{\min} - k_2 y_{\min} = 0.78 = \varepsilon$，此时有 $T_1 = T_2 = 4.78$。观察第 2 组，y_3 的度量值变为 5，第 3 组数据，y_3 权重变为 0.35 时，原模型中控制因子 ε 仍然是 0.78，而改进后的模型中控制因子由于是 $y_{\min}$ 与 $\alpha_{\min}$ 的函数，分别变为 0.81 与 0.94，这样的度量结果相比较原模型，更加强调了最小关键属性的度量值与权重对整个软件可信性的影响，由表 2 第 2、3 行可以看出，T_2 对于 T_1 有了不同程度的下降，即从 4.61 和 5.56 分别降到 4.56 和 5.2，这是符合软件实际情况的，因此，我们的度量模型更具有优越性。

5. 结束语

本文通过对文献[1]中提出的软件可信性度量模型的深入分析，针对原模型中与软件实际情况不相符的事实，提出了一种改进的基于属性划分的数学模型，该模型强调了最小度量值的关键属性及其权重对整个软件可信性的影响，对比原模型，更加符合实际情况，具有一定的优越性。下一步的工作主要侧重于如何确定区分关键属性与非关键属性以及各个属性权重的确定。

参考文献 (References)

[1] H. W. Tao, Y. X. Chen. A metric model for trustworthiness of softwares. 2009 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology, 15-18 September 2009: 69-72.

[2] J. P. Anderson. Computer security technology planning study. ESD-TR-73-51. Vol. 1, AD-758 206, ESD/AFSC, Hanscom AFB, Bedford, 1972.

[3] ISO/IEC. Information technology-security techniques-evaluation criteria for IT security. Part 1: Introduction and General Model, 2005.

[4] A. Algirdas, J. C. Laprie, R. Brian, et al. Basic concepts and taxonomy of dependable and secure computing. IEEE Transactions on Dependable Secure, 2004, 1(1): 11-33.

[5] 王怀民, 唐扬斌, 尹刚等. 互联网软件的可信机理[J]. 中国科学: E 辑, 2006, 36(10): 1156-1169.

