doi:10.11887/j.cn.201703011 http://journal.nudt.edu.cn

微处理器容软错误设计量化评估指标及评估方法*

龚 锐,郭御风,邓 宇,石 伟,窦 强
(国防科技大学 计算机学院,湖南 长沙 410073)

摘 要:针对高可靠微处理器软容错设计,提出了一种新的可靠性度量标准,增强的平均无失效工作量,以解决现有度量标准没有综合考虑性能、面积、功耗开销带来的可靠性降低的缺点;提出了一种评估方法对增强的平均无失效工作量以及两种控制流检测技术进行定量评估。评估结果表明,软硬件结合的控制流检测技术较好地折中了可靠性、性能、面积和功耗。量化评估指标全面考虑了多种开销对微处理器可靠性的影响,采用相应的评估方法可以更加准确地对微处理器可靠性加固手段进行定量评估,以指导设计探索和设计优化。

关键词:容软错误;量化评估;评估方法;微处理器;可靠性
中图分类号:TP302.8 **文献标志码**:A **文章编号**:1001-2486(2017)03-064-05

Quantitative evaluation metric and methodology for microprocessor soft error tolerance design

GONG Rui, *GUO Yufeng*, *DENG Yu*, *SHI Wei*, *DOU Qiang*
(College of Computer, National University of Defense Technology, Changsha 410073, China)

Abstract: Aiming at highly reliable microprocessor soft error tolerance design, a new metric, eMWTF (enhanced mean work to failure), was proposed to capture the trade-off among reliability, performance, area and power. A quantitative approach for evaluating eMWTF was also presented. Two control flow checking techniques were quantitatively evaluated in reliability. The experimental results indicate that the control flow checking by compiler signatures and hardware checking achieves better trade-off among reliability, performance, area and power. Because the eMWTF metric takes into consideration performance, area and power overheads, the quantitative reliability evaluation can be more accurate by using this metric and corresponding methodology. Finally, the evaluation results can effectively guild the design exploring and optimization.

Key words: soft error tolerance; quantitative evaluation; evaluation methodology; microprocessor; reliability

应用于复杂电磁环境的集成电路受到高能粒子轰击,会发生瞬时充放电,使得逻辑状态发生翻转,这种由高能粒子轰击所引发的错误被称为“软错误”。高可靠微处理器一般采用多种容软错误设计技术。这些容软错误设计在提高微处理器可靠性的同时,不可避免地带来了性能、面积、功耗的开销。最新的软错误发生机理研究表明,性能、面积、功耗的开销对于微处理器的可靠性有负面影响。

1 研究背景

1.1 软错误类型

高能粒子引起的微处理器软错误包括单事件翻转(Single Event Upset, SEU)、单事件瞬态(Single Event Transient, SET)和多位翻转(Multi Bit Upsets, MBU)等。其中SEU是指单个存储单元遭到高能粒子轰击而发生的逻辑翻转。翻转后错误的值将一直被保持到下一次写入操作。SET是指高能粒子轰击导致组合逻辑通路上产生的毛刺。这种SET毛刺有可能沿组合逻辑通路传递,也可能被电路自身的结构所屏蔽。当SET毛刺恰好在时钟沿传递到时序逻辑输入,错误的值将会被采样,导致微处理器功能错误。此外,随着集成电路特征尺寸的缩小和集成度的提高,一次粒子轰击有可能导致动态随机访问存储器(Dynamic Random Access Memory, DRAM)存储阵列或静态随机访问存储器(Static Random Access Memory, SRAM)存储阵列内相邻的多个存储单元发生翻转,这种类型的软错误被称为MBU。

* 收稿日期:2015-11-13
基金项目:国家自然科学基金资助项目(61202123,61202122,61402497)
作者简介:龚锐(1980—),男,四川雅安人,助理研究员,博士,E-mail:rgong@nudt.edu.cn

与设计制造过程中引入的硬错误相比,上述软错误具有瞬态、可恢复、发生位置和时间随机等特点。

1.2　软错误发生机理

电子器件发生软错误的概率受辐射水平、存储电荷及敏感源漏区域面积的影响。一般采用软错误率(Soft Error Rate,SER)[1]来表征器件发生软错误的概率。SER 可以采用式(1)推算[2]。

$$SER \propto F \cdot A_{sd} \cdot \exp\left(-\frac{Q_{crit}}{Q_s}\right) \tag{1}$$

其中:F 是能量大于 1 MeV 的高能粒子流密度;A_{sd}是对辐射敏感的面积,对单个晶体管器件来说,即源漏极面积;Q_{crit}是导致芯片中存储信息发生逻辑翻转所需要的最小电量,称为临界电量[3-4];Q_s 则是粒子轰击在芯片上引起的实际充放电量。

1.3　容软错误能力量化评估

一般来说,对软错误进行检测、屏蔽与恢复,都需要某种冗余机制。这些冗余设计不可避免地带来了芯片面积、程序执行性能和微处理器功耗的开销。微处理器受到粒子轰击的概率正比于其暴露于辐射环境中的芯片面积,芯片面积的增加将导致更多的软错误。程序执行性能的降低,将增加单个程序的执行时间,从而增加单个程序执行过程中受到高能粒子轰击的概率。微处理器功耗的上升将导致芯片工作温度的升高,根据国内外研究人员在电路级的研究,SEU 对温度的变化不敏感[5],但在 -55 ~ +125 ℃范围内,SET 毛刺的宽度随温度的升高而变大,其宽度与温度基本呈线性变化[6]。SET 毛刺的展宽将增加其被下级时序逻辑单元采样到的概率,从而增加微处理器发生软错误的概率。因此,冗余设计带来的面积、性能、功耗的开销对微处理器的容软错误能力是有负面影响的。片面强调容软错误的冗余设计而忽视其开销带来的负面影响,并不一定能获得最优化的可靠性提升。

2　相关工作

可靠性评估中重要的量化评估指标是平均无失效时间(Mean Time To Failure,MTTF),该参数表示微处理器发生失效的期望时间。在容软错误能力评估中,MTTF 可以简单表示为:

$$MTTF = \frac{1}{SER} \tag{2}$$

由于相当一部分的软错误会被微处理器体系结构的固有特性或各种软错误加固技术所屏蔽,并不会引起程序的执行结果发生错误。因此,文献[7]提出了结构弱点因子(Architectural Vulnerability Factor,AVF)来表示原始软错误导致微处理器失效的概率,该参数也可以表征微处理器体系结构所具有的软错误屏蔽能力。在此基础上 MTTF 可以更精确地表示为:

$$MTTF = \frac{1}{SER \cdot AVF} \tag{3}$$

采用 MTTF 进行可靠性评估,只考虑了容软错误技术带来的可靠性提升(即 AVF 的降低),而未考虑其面积、性能、功耗开销带来的可靠性降低。文献[8]给出了平均无失效指令(Mean Instruction To Failure,MITF)的概念。MITF 表征微处理器在失效前可以执行的平均指令条数,可以表示为:

$$MITF = IPC \cdot Frequency \cdot MTTF = \frac{IPC \cdot Frequency}{SER \cdot AVF} \tag{4}$$

其中,IPC 表示每周期可执行的指令条数,$Frequency$ 表示微处理器频率。

文献[9]进一步推广,提出了平均无失效工作量(Mean Work To Failure,MWTF)的概念来表征微处理器在失效前可完成的平均工作量。MWTF 定义为:

$$MWTF = \frac{1}{SER \cdot AVF \cdot t_{exe}} \tag{5}$$

其中,t_{exe}为微处理器执行给定工作所需的时间,一般表示为执行一组测试程序所需的时间。

MITF 和 MWTF 两个量化指标考虑了性能开销对微处理器容软错误能力的影响,但仍未考虑面积和功耗的影响。在前期的研究中,提出了改进的平均无失效工作量(modified MWTF,mMWTF)的概念[10],将面积和性能开销都纳入量化评估指标内。mMWTF 定义为:

$$mMWTF = \frac{1}{SER \cdot A \cdot AVF \cdot t_{exe}} \tag{6}$$

其中 A 表示芯片面积。

上述相关工作一步步推进可靠性量化评估向更全面的方向发展,但仍未将功耗因素考虑在内。

3　量化评估指标

已有的可靠性量化评估指标中,一般采用 SER 来表征微处理器在单位时间内发生的软错误。由于芯片面积不同,辐射面积就不相同,SER 也不同。精确定义瞬态故障率(Transient Fault Rate,TFR)为单位芯片面积微处理器在单位时间

内发生 SEU、SET 等瞬态故障的概率。可以认为，在相同的制造工艺和相同的辐射条件下，微处理器的 TFR 相同。

现只考虑 SEU 和 SET 两种类型的瞬态故障。定义 AVF 为 SEU 导致微处理器发生失效的概率。定义 TVF 为 SET 被寄存器采样而发生 SEU 的概率。AVF 表征了体系结构和软件对 SEU 的屏蔽能力，而 TVF 则表征了寄存器采样窗口对 SET 的屏蔽能力。SET 被采样形成 SEU 后，也只有 AVF 导致微处理器发生失效。假设微处理器中发生的 SET 占所有瞬态故障的百分比为 P_{SET}，则在单位时间内微处理器发生失效的次数为：

$$\begin{aligned} N_{\mathrm{f}} &= N_{\mathrm{f\_SET}} + N_{\mathrm{f\_SEU}} \\ &= TFR \cdot A \cdot P_{\mathrm{SET}} \cdot TVF \cdot AVF + TFR \cdot A \cdot (1 - P_{\mathrm{SET}}) \cdot AVF \\ &= TFR \cdot A \cdot AVF \cdot [P_{\mathrm{SET}} \cdot TVF + (1 - P_{\mathrm{SET}})] \end{aligned} \tag{7}$$

即单位时间内发失效的次数 N_{f} 为由 SET 引起的失效数 $N_{\mathrm{f\_SET}}$ 和由 SEU 引起的失效数 $N_{\mathrm{f\_SEU}}$ 的总和。

对于 SET 来说，其 TVF 可以近似表征为 SET 脉冲宽度 W 与寄存器时钟频率 T_{clk} 的比值，即：

$$TVF = \frac{W}{T_{\mathrm{clk}}} \tag{8}$$

由国内外对软错误机理的研究可知，温度 T 对 SEU 基本没有影响，但会导致 SET 脉冲宽度 W 展宽，且 W 与 T 基本呈线性关系。假设 W 与 T 的关系为：

$$W = aT + b \tag{9}$$

假设 SET 脉冲展宽后仍然小于等于时钟频率 T_{clk}，那么将式(9)代入式(8)，有：

$$TVF = \frac{aT + b}{T_{\mathrm{clk}}} \tag{10}$$

可以简单地认为微处理器工作温度与单位面积功耗(P/A)即功耗密度呈线性关系，所以有：

$$T = x\frac{P}{A} + y \tag{11}$$

将式(11)代入式(10)，可得：

$$TVF = \frac{a\left(x\frac{P}{A} + y\right) + b}{T_{\mathrm{clk}}} = \frac{\alpha\frac{P}{A} + \beta}{T_{\mathrm{clk}}} \tag{12}$$

即在时钟频率不变的情况下，由 SET 导致 SEU 的概率 TVF 与单位面积功耗(P/A)呈线性关系。将式(12)代入式(7)，有：

$$N_{\mathrm{f}} = TFR \cdot A \cdot AVF \cdot \left[P_{\mathrm{SET}} \cdot \frac{\alpha\frac{P}{A} + \beta}{T_{\mathrm{clk}}} + (1 - P_{\mathrm{SET}})\right] \tag{13}$$

因此，提出增强的平均无失效工作量(enhanced Mean Work To Failure，eMWTF)，来表征微处理器在发生失效前可以完成的平均工作量。该量化标准可定义为：

$$eMWTF = \frac{1}{N_{\mathrm{f}} \cdot t_{\mathrm{exe}}} \tag{14}$$

其中，t_{exe} 为完成单位工作量所需的时间，一般表征为完成一组典型测试程序所需的时间。因此 eMWTF 可以表示微处理器在失效前可以完成这种典型测试程序的次数，即可以完成的平均工作量。将式(13)代入式(14)，可得：

$$eMWTF = \frac{1}{TFR \cdot A \cdot AVF \cdot \left[P_{\mathrm{SET}} \cdot \frac{\alpha\frac{P}{A} + \beta}{T_{\mathrm{clk}}} + (1 - P_{\mathrm{SET}})\right] \cdot t_{\mathrm{exe}}} \tag{15}$$

由式(15)可知，eMWTF 是一个涉及了多种设计维度的微处理器可靠性量化评估指标，该指标综合考虑了软错误发生的机理(TFR 和 P_{SET})、容软错误设计带来的可靠性的提升(即 AVF 的降低)以及容软错误设计带来的面积(A)、性能(T_{clk} 和 t_{exe})和功耗(P)开销对可靠性的影响。因此是一个全面准确的量化评估指标。

4 量化评估方法

针对 eMWTF 的量化评估方法如图 1 所示。

该量化评估方法紧密结合半定制的微处理器设计流程。首先在 RTL 级的功能模拟时，执行一组标准的测试程序，获得 t_{exe} 参数。

功能模拟通过后，由综合工具将 RTL 级代码综合为门级网表。在综合的过程中可以知道该设计可以运行的时钟频率 T_{clk} 和所使用的标准单元总面积。由于芯片的总面积 A 需要在最后版图生成后才能确认，但量化评估需要尽可能地在设计的早期进行，所以采用综合时获得的标准单元和 SRAM 总面积信息来近似替代芯片总面积 A。

通过综合得到门级网表以后，需要在门级网表上进行错误注入模拟，以获得微处理器的 AVF 参数。同时，可以利用前仿时得到的波形信息，通过功耗评估工具获得较准确的功耗参数 P。

在获得 t_{exe}、A、T_{clk}、AVF、P 等参数以后，可以对 eMWTF 进行量化评估，从而指导设计折中和设计选择。如果没有达到预设的可靠性指标，则需要迭代回去重新进行设计。

需要注意的是，由于 eMWTF 中的某些参数并不能获得准确的数值。如式(15)中的瞬态故障率 TFR 与使用环境的辐照水平相关；SET 故

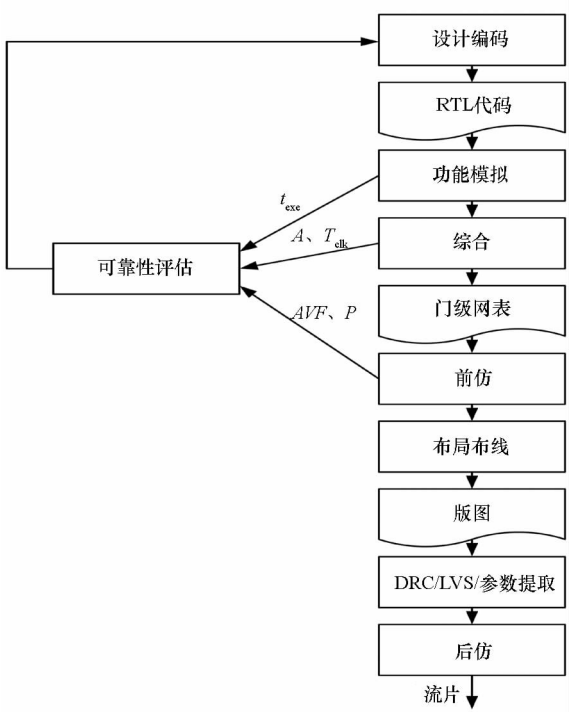

图 1　量化评估方法

Fig. 1　Quantitative evaluation methodology

障占所有瞬态故障的百分比 P_{SET}除了与辐照水平相关,还与芯片内部逻辑相关;TVF 与单位面积功耗(P/A)的线性关系参数 α、β 则与辐照水平、工艺、电气特性等相关。在评估时,只能对这些参数进行假设。此外,在评估的过程中也进行了一些假设,如采用综合时获得的标准单元和 SRAM 总面积来近似替代芯片总面积 A。因此,采用 eMWTF 和所提出的评估方法,无法获得精确的可靠性数值。但是可以评估出在相同假设条件下,不同可靠性加固手段所能带来的相对的可靠性关系,从而指导设计空间探索和设计选择。

5　量化评估实验及结果

采用 eMWTF 指标对两种控制流检测技术进行可靠性量化评估,并给出相应的结果。

5.1　控制流检测技术

高能粒子导致的故障可能引起控制流错误,即程序的执行流程发生混乱。控制流检测的基本思想是实时监测程序的运行轨迹并与编译预期的轨迹进行比较,以有效防止由于控制流错误导致的系统崩溃。

5.1.1　CFCSS 技术

文献[11]提出了一种纯软件实现的控制流检测(Control Flow Checking by Software Signatures,CFCSS)技术。该方法定义程序流图为有向图 $CFG = (V,E)$,其中 $V = \{v \mid v$为基本块$\}$,$E = \{\langle v_i, v_j\rangle \mid$ 存在从 v_i 到 v_j 的分支或跳转$\}$。对于某个特定的基本块 v_i,赋予其唯一的签名值 S_i。如果 $\exists \langle v_i, v_j\rangle \in E$,则 v_i 到 v_j 的签名距离 $d_j = S_i \oplus S_j$,该签名距离在编译时即可确定。当程序执行从 v_i 到 v_j 的控制流转移时,计算运行时签名值 $s_j = S_i \oplus d_j$。如果分支或转移正确,则 $s_j = S_i \oplus d_j = S_i \oplus (S_i \oplus S_j) = S_j$。如果 $s_j \neq S_j$,则表明发生了控制流错误。由于采用纯软件实现,CFCSS 比较灵活,且不用对硬件进行任何改动,没有额外的面积开销。但是这些签名检测指令若编译为 8051 指令,执行一次签名检测需要 13 个时钟周期,性能开销比较大。

5.1.2　CFCCH 方法

为了解决 CFCSS 技术性能开销大的缺点,文献[12]中提出了一种编译签名硬件检测的控制流检测(Control Flow Checking by Compiler signatures and Hardware checking, CFCCH)方法。该方法采用与 CFCSS 技术相同的签名算法,但只在每个基本块的头部依次插入三个字节的签名数据,即签名距离 d_i、签名值 S_i 和运行时调整签名 D_i。为了实现硬件检测,增加了两个特殊寄存器 Sreg 和 Dreg,分别记录当前基本块的签名值并运行时调整签名。在每次控制流转移,即分支或跳转指令之后,硬件自动进行一次检测,若检测无误,才运行新基本块的指令。每次检测只需要 3 个时钟周期。CFCCH 方法采用硬件进行检测,有额外的面积开销,但是性能开销大大减少。

5.2　评估结果

上述 CFCSS 和 CFCCH 两种控制流检测技术各有优劣。分别采用 MTTF、MWTF、mMWTF 和 eMWFT 4 种量化评估指标对这两种容软错误技术进行评估,并且对未经加固设计的 8051 也进行量化评估,以获得两种加固技术相对于未加固芯片的归一化可靠性参数,从而指导设计选择。

循环运行测试程序集,并注入了 10 000 个故障,以使结果具有统计意义。同时,采用65 nm工艺对 3 款微控制器进行了综合,约束的时钟频率均为 100 MHz。从综合得到的总的标准单元和 SRAM 面积来看,CFCSS 由于没有任何硬件改动,没有额外的面积开销。CFCCH 采用了硬件检测,总的标准单元面积比未采用容软错误技术的非容错(NOn Fault Tolerance, NFOT)版本大 7.5%。此外,采用功耗评估工具对 3 款微控制器进行功耗评估。结果表明,CFCSS 与 NOFT 功耗相当,但 CFCCH 比 NOFT 约增加了 10.3% 的功耗。运行

了一组测试程序,以便获得性能参数。结果表明,CFCSS 的签名检测代码执行一次需要 13 个周期,性能开销较大,其性能开销为 NOFT 的 54% ~112%。CFCCH 在每个检测点只增加了额外的 3 个时钟周期,带来了 9% ~36% 的性能开销,低于 CFCSS。

在获得上述 *AVF*、面积、功耗、性能参数的基础上,为了获得 eMWTF 数值,做出如下假设。假设 P_{SET} 为 0.5,即发生 SET 故障的概率和 SEU 故障概率相同。假设线性关系参数 $\alpha=\beta=1$。在上述假设基础上,获得的评估结果如图 2 所示。

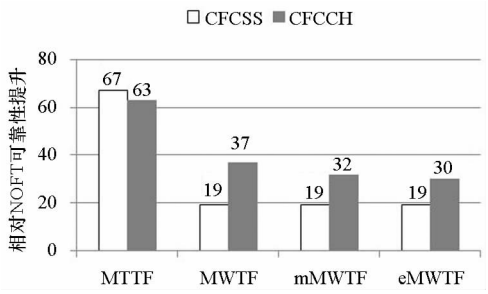

图 2 归一化可靠性评估结果

Fig. 2 Normalized reliability evaluation results

图 2 中的可靠性数值都是与 NOFT 进行了归一化后的相对值。从图 2 中可以看出,由于具有性能开销,同一容软错误技术的 MWTF 值要小于其 MTTF 值。MWTF 表示两次失效之间能够执行的平均工作量,而工作量的定义与实际应用相关。这说明对实际的应用来说,尽管容软错误技术使得两次失效之间能够正常执行的时间大大增加,但是这段时间内所能执行的有用工作量并没有成比例增加。同样地,具有面积开销的容软错误技术,其 mMWTF 值也要小于 MWTF 值。这是因为面积的开销将导致更多的原始软错误。此外,具有功耗开销的容软错误技术,其 eMWTF 值也要小于 mMWTF 值,这是因为功耗开销将导致芯片温度的上升,从而引起 SET 脉冲宽度变大,使其更容易引发寄存器翻转。从以上分析可知,对容软错误技术进行评估时必须全面、定量地考虑性能、面积和功耗的开销,以便进行更好的折中。

6 结论

为解决原有微处理器容软错误评估中不考虑功耗开销的缺点,本文提出了一种新的可靠性度量标准 eMWTF。该标准全面考虑性能、面积、功耗开销对可靠性带来的负面影响。与传统的度量标准相比,eMWTF 能够更加准确地定量表征微处理器的可靠性,因而更具指导意义,能够有效地指导设计探索和选择。

参考文献(References)

[1] Ziegler J F, Curtis H W, Muhlfeld H P, et al. IBM experiments in soft fails in computer electronics (1978—1994)[J]. IBM Journal of Research and Development, 1996, 40(1): 3 - 18.

[2] Mukherjee S. Architecture design for soft errors[M]. UK: Morgan Kaufmann Press, 2008.

[3] Tang H H K. Nuclear physics of cosmic ray interaction with semiconductor materials: particle-induced soft errors from a physicist's perspective[J]. IBM Journal of Research and Development, 1996, 40(1): 91 - 108.

[4] Freeman L B. Critical charge calculations for a bipolar SRAM array[J]. IBM Journal of Research and Development, 1996, 40(1): 119 - 129.

[5] Truyen D, Boch J, Sagnes B, et al. Temperature effect on heavy-ion induced parasitic current on SRAM by device simulation: effect on SEU sensitivity[J]. IEEE Transactions on Nuclear Science, 2007, 54(4): 1025 - 1029.

[6] 梁斌, 陈书明, 刘必慰. 温度对数字电路中单粒子瞬态脉冲的影响[J]. 半导体学报, 2008, 29(7): 1407 - 1411.
LIANG Bin, CHEN Shuming, LIU Biwei. Temperature dependence of digital single event transient[J]. Journal of Semiconductors, 2008, 29(7): 1407 - 1411. (in Chinese)

[7] Mukherjee S S, Weaver C, Emer J, et al. A systematic methodology to compute the architectural vulnerability factors for a high-performance microprocessor[C]//Proceedings of IEEE/ACM International Symposium on Microarchitecture, 2003: 29 - 40.

[8] Weaver C, Emer J, Mukherjee S S, et al. Techniques to reduce the soft error rate of a high-performance microprocessor[C]// Proceedings of International Symposium on Computer Architecture, 2004: 264 - 275.

[9] Reis G A, Chang J, Vachharajani N, et al. Design and evaluation of hybrid fault-detection systems[C]//Proceedings of International Symposium on Computer Architecture , 2005: 148 - 159.

[10] Gong R, Dai K, Wang Z Y. A framework to evaluate the trade-off among AVF, performance and area of soft error tolerant microprocessors [C]//Proceedings of IEEE International Symposium on Defect and Fault Tolerance in VLSI Systems, 2008: 184 - 192.

[11] Oh N, Shirvani P P, McCluskey E J. Control flow checking by software signatures[J]. IEEE Transactions on Reliability, 2002, 51(1): 111 - 122.

[12] 龚锐, 陈微, 刘芳, 等. 一种软硬件结合的控制流检测与恢复方法[J]. 计算机研究与发展, 2009, 46(2): 345 - 351.
GONG Rui, CHEN Wei, LIU Fang, et al. Control flow checking and recovering by compiler signatures and hardware checking [J]. Journal of Computer Research and Development, 2009, 46(2): 345 - 351. (in Chinese)