

东软机密

Neusoft

文件编号: D00-CMG015

测试估计指南

版本: 1.1.0-0.0.0

2009-6-12

东软集团股份有限公司 过程改善中心

(版权所有, 翻版必究)

文件修改控制

[illegible]

目录

1. 目的.....	1
2. 估计的步骤.....	1
2.1 任务分解（WBS）	1
2.2 确定估计方法和参数.....	2
2.3 规模估计	2
2.4 工作量估计	3
2.5 成本估计	4
2.6 进度估计	5
2.7 质量估计	5
3. 附录一：规模估计的方法.....	7
3.1 WIDEBAND DELPHI TECHNIQUE.....	7
3.2 PERT SIZING TECHNIQUE.....	9
3.3 类比法（ANALOGY APPROACH）	10
3.4 运算法（ALGORITHMIC APPROACH）	11
3.4.1 基于代码规模的运算法.....	11
3.4.2 基于需求抽样的运算法.....	11
4. 附录二：工作量估计的方法.....	13
4.1 运算法（ALGORITHMIC APPROACH）	13
4.2 专家判断法（EXPERT JUDGEMENT APPROACH）	13

1. 目的

本文描述的是测试项目估计活动的步骤和方法，根据本文可以估计测试项目的规模、工作量、成本、进度和质量。

2. 估计的步骤

在测试项目生命周期开始时，必须对将要测试项目进行估计，并且对项目的估计、跟踪和优化等应该贯穿整个生命周期。步骤如下：

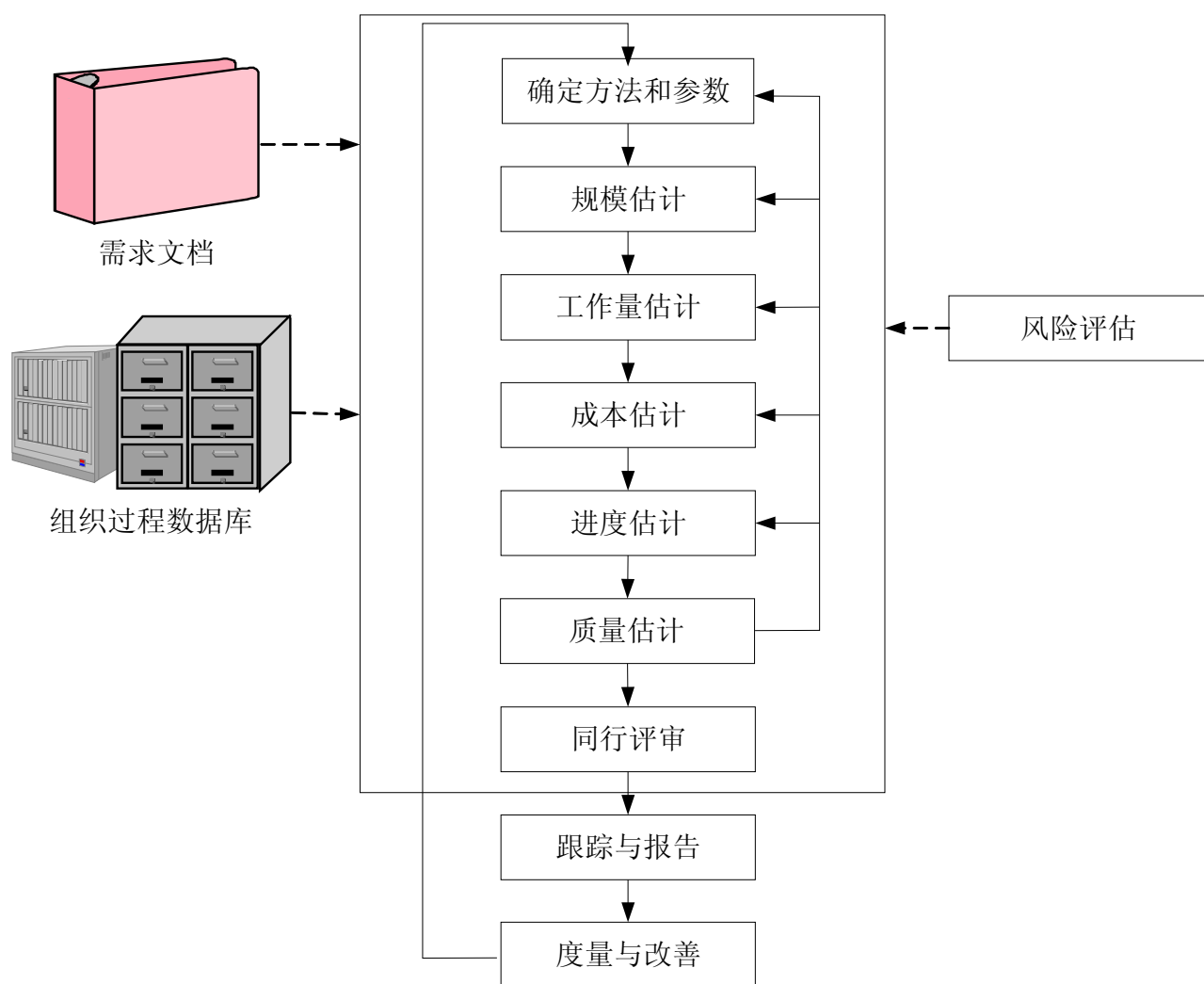


图 2-1 项目估计步骤

2.1 任务分解（WBS）

1. 面向产品的分解结果是确认项目范围、进行项目估计的基础，以一个合理的方式把产品分解成一个由多级父节点和子节点构成的树状的层次结构。任务分解的一般规则包

括：

- WBS 的树结构遵从树的一般定义，每个父节点可以有多个子节点，而每个子节点最多只有一个父节点。
- WBS 的叶子节点的规模是可估计的，分解结果的层次不是无限的扩展，如果某一级的节点规模已经可以较为准确的估计，那么下级节点可以从 WBS 中摘除。

2. 任务分解树的各级节点一般有以下含义：

- Level1: 通常直接标识整个系统或项目
- Level2: 系统中的子系统。
- Level3: 一般指子系统模块。
- Level4: 一般指模块中的各个子模块。

3. 将任务分解结果记录到“测试计划”中。

2.2 确定估计方法和参数

1. 根据已知的项目范围、业务类型、技术路线、项目生命周期模型等特征，查询组织 PCB，选择适当版本的 PCB 数据用于后续的估计。
2. 当组织没有 PCB 数据时，在组织的过程数据库中查找可以参考的历史数据用于后续的估计。
3. 根据项目的特征和已经选定的 PCB 或者历史项目数据，确定在估计过程中使用的估计方法和参数。必要时，对这些参数进行适当调整。如果估计的方法和参考的历史数据之间存在较大的差异或对参数进行了较大的调整，应当给出适当的说明。

2.3 规模估计

测试项目规模通常使用测试用例数（条）为单位进行估计，也可以根据项目实际情况使用功能点、代码行、需求数等方式来进行规模估计。规模估计是工作量、质量和进度估计的基础，也是后续估计活动的必要输入。估计测试规模主要依靠历史数据和相似系统的经验以及合适的估计方法，因此参与估计的人应当具有相似系统的经验，结合软件过程数据库中积累的历史数据，使用合适的估计方法，综合进行测试规模的估计。

使用测试用例为单位进行测试规模估计可以按下面的步骤进行：

1. 估计重用测试用例规模

估计重用测试用例时需要考虑，与重用测试用例对应的测试数据是否能够使用。重用的测试用例有哪些潜在的风险，重用的测试用例需要进行修改的比率有多大。总之对重

用测试用例需要进行非常慎重的选取和评估。

2. 估计新开发测试用例规模

在已知被测试项目代码行规模的前提下，使用附录一中的 3.4.1 基于代码规模的运算法，对任务分解结果中的不可继续分解的子节点进行测试用例规模估计；在已知被测试项目需求规模的前提下，使用附录一中的 3.4.2 基于需求规模的运算法，对项目测试用例规模进行估计。参考选定的历史数据和项目的特征对估计结果进行适当调整，对于那些进行了较大调整的估计结果进行说明，并记录到“测试计划”中。项目也可以根据实际情况，使用附录一中其他规模估计方法对项目测试用例规模进行估计。

3. 估计测试用例总规模

估计重用测试用例规模加上新开发测试用例规模，就是估计的测试用例总规模。

4. 将测试用例规模估计结果记录到“测试计划”

2.4 工作量估计

工作量估计是使用一个适当的估计关系将项目规模估计转化为工作量估计。在测试项目中，通常应当包含以下内容：

1. 需求理解的工作量
2. 大纲设计的工作量
3. 用例设计的工作量
4. 测试执行的工作量
5. 项目管理的工作量
6. 质量保证的工作量
7. 配置管理的工作量
8. 培训的工作量

测试项目的工作量估计可以按照下面的步骤进行：

1. 计算用例设计的工作量

测试用例设计的工作量包括测试用例设计工作量和测试数据设计工作量。

根据测试项目新开发测试用例规模和组织的测试用例设计效率，计算测试用例设计的工作量，并根据项目的人员能力进行适当的调整，结果记录到“测试计划”中。具体方法参考附录二中的运算法。

2. 计算测试执行的工作量

测试执行的工作量包括测试环境搭建工作量、接受测试工作量、测试用例执行工作量

和缺陷管理工作量。

根据测试项目总测试用例规模、测试执行轮次以及测试用例执行效率，计算测试执行的工作量。具体方法参考附录二中的运算法

3. 计算需求理解的工作量

根据估计的测试用例设计工作量、历史项目数据中工作量分布，计算需求理解的工作量。

4. 计算大纲设计的工作量

根据估计的测试用例设计工作量、历史项目数据中大纲设计工作量分布，计算大纲设计的工作量。

5. 计算项目管理的工作量

根据测试设计阶段（包括需求理解、大纲设计、用例设计、测试执行）工作量、历史项目数据中项目管理工作量分布，计算项目管理的工作量。

6. 计算质量保证的工作量

根据测试设计阶段工作量、历史项目数据中质量保证工作量分布，计算质量保证的工作量。

7. 计算配置管理的工作量

根据测试设计阶段工作量、历史项目数据中配置管理工作量分布，计算配置管理的工作量。

8. 计算培训的工作量

根据测试设计阶段工作量、历史项目数据中培训工作量分布，计算培训的工作量。

2.5 成本估计

成本估计通常应包含开发费用、设备费用、差旅费用、培训等其他费用的估计。成本估计可以按照下面的步骤进行：

1. 计算开发费用

根据项目的特点及各项开发活动需要确定项目所需各类资源、资源的单价、及其投入的工时来计算开发费用。

2. 计算设备费用

根据项目的特点及需要确定项目所需各类设备、设备的单价、数量来计算设备费用。

3. 估计差旅费用

根据项目的特点及需要确定项目所需要出差的地点、人数、周期，并参考历史项目数据和当时当地的交通费用、住宿费用、餐饮费用等估计差旅费用。

4. 估计其他费用

根据项目的特点及需要确定项目所需其他费用，并参考历史项目数据估计其费用。

2.6 进度估计

在正常情况下，项目的完成时间和主要里程碑是根据项目的工作量和可分配资源的数量来决定，但有时候项目中主要里程碑和完成的时间就已经有了明确的规定，这样就要通过调整可分配的资源来满足项目的里程碑和完成时间的要求，这项工作必须由经验丰富的人来完成。对进度的估计可以按下面的步骤进行：

1. 将本项目所有的可用资源进行分类。
2. 分析项目各个任务之间的依赖关系，确定任务执行的顺序，协调可以并发的任务，确保他们能在后续任务需要其工作成果之前完成。
3. 根据任务的依赖关系和执行顺序确定关键路径。
4. 在关键路径上定义里程碑，根据关键路径上的任务和可以分配的资源，确定出里程碑的时间和完成时间。
5. 将进度估计的结果记录到“测试计划”中。

2.7 质量估计

对于测试项目来说，包括软件质量和测试质量两个部分。

软件质量是对被测试工作产品或服务的评价，测试质量是对测试过程的评价，软件质量是产品或服务的一种属性，测试质量是过程的一种属性，通常情况下，通过不断的改进测试过程的质量来提升软件产品或服务的质量。

软件质量估计可以按照下面的步骤进行：

1. 确定被测试模块或子模块规模和缺陷密度。
2. 依据模块或子模块规模和缺陷密度，计算各模块或子模块的缺陷数。
3. 测试负责人将估计的缺陷数分配到测试轮次中。
4. 测试负责人根据各模块或子模块缺陷估计的结果，与测试目标进行比较，判断目标是否能够达成，如果目标不能实现，测试负责人与模块负责人共同协商调整模块缺陷数。

5. 将软件质量估计的结果记录到“测试计划”中。

测试质量在测试项目中主要是指需求理解、大纲设计、用例设计、测试执行阶段/活动缺陷情况，测试质量估计可以按照下面的步骤进行：

1. 设定各阶段/活动的缺陷密度

根据组织能力标准或项目目标要求设定各阶段/活动的缺陷密度。

2. 计算项目级活动需求理解/大纲设计/测试用例设计测试执行的缺陷数

需求理解/大纲设计/用例执行/测试执行等这些项目级活动的缺陷数可以根据各阶段/活动输出的成果物规模和各阶段/活动的缺陷密度计算这些活动的缺陷数。

3. 计算出项目活动总缺陷数

将项目及活动的缺陷数进行汇总，得出项目活动总缺陷数。

4. 将测试质量估计结果记录到“测试计划中”

3. 附录一：规模估计的方法

3.1 Wideband Delphi Technique

Wideband Delphi Technique 是一种鼓励参加估计的人员之间就相关问题进行讨论的方法，用这种方法能充分发挥集体的力量，使估计的结果更切合实际，下面的表格是使用这样方法估计的步骤：

步骤	活动
1.	召集人召集所有参加估计的人员，并将项目的需求和估计用表格分发给大家
2.	召集人召集所有参加估计的人员进行一个会议，讨论有关测试用例规模的问题
3.	参加估计的每个人匿名的填写估计表格
4.	召集人收集所有的估计表格，然后形成反馈表返回给参加估计的人员
5.	召集人召集所有参加估计的人员进行一个会议，主要是讨论估计上差异
6.	参加估计的人员根据讨论的结果，在反馈表上提交另一个匿名的估计
7.	重复 4~6 直到达成关于测试用例规模最大程度的一致

下面是估计反馈表的例子：

项目名称:	TL:
测试策略:	测试工具:
这是第_轮的估计范围:	下次:
模块A:	
0 20 40 60 80 100 120	____条(测试用例)
模块B:	
0 20 40 60 80 100 120	____条(测试用例)
模块C:	
0 20 40 60 80 100 120	____条(测试用例)
模块D:	
0 20 40 60 80 100 120	____条(测试用例)
模块E:	
0 20 40 60 80 100 120	____条(测试用例)
模块F:	
0 20 40 60 80 100 120	____条(测试用例)
模块G:	
0 20 40 60 80 100 120	____条(测试用例)
X – 估计	
X* – 你的估计	
XI – 估计的中间值	
备注:	
[在这里写下需要特殊说明的问题]	

举例说明如下:

假如由一个由 A、B 两个模块组成的项目: 例子项目, 甲是召集人, 乙和丙是参加估计的人。

首先, 甲将项目相关的资料和估计表格发给乙和丙, 并召开一个甲乙丙参加的会议, 分析项目的问题。

乙估计 A 模块用 40 条测试用例, B 模块用 80 条测试用例, 丙估计 A 模块用 80 条测试用例, B 模块用 120 条测试用例, 然后乙和丙分别匿名的填写估计表格, 甲收集表格后, 进行汇总, 然后将下面表格反馈给乙和丙。

项目名称:

TL:

测试策略:

测试工具:

这是第 1 轮的估计范围:

下次:

范围:

模块A:

		40(X)	60(XI)	80(X)			
0	20	40	60	80	100	120	条(测试用例)

模块B:

				80(X)	100(XI)	120(X)	
0	20	40	60	80	100	120	条(测试用例)

模块C:

0	20	40	60	80	100	120	条(测试用例)
---	----	----	----	----	-----	-----	---------

模块D:

0	20	40	60	80	100	120	条(测试用例)
---	----	----	----	----	-----	-----	---------

模块E:

0	20	40	60	80	100	120	条(测试用例)
---	----	----	----	----	-----	-----	---------

模块F:

0	20	40	60	80	100	120	条(测试用例)
---	----	----	----	----	-----	-----	---------

模块G:

0	20	40	60	80	100	120	条(测试用例)
---	----	----	----	----	-----	-----	---------

X – 估计

X* – 你的估计

XI – 估计的中间值

备注:

[在这里写下需要特殊说明的问题]

甲召集乙和丙举行一个会议讨论估计的差异，然后分别重新估计，填写估计表，甲收集后重新汇总，形成新的反馈表交给大家，重复这个过程，直到乙和丙在规模上达成一致，完成对测试用例规模的估计。

3.2 Pert Sizing Technique

这种方法共估计三个值：测试用例规模的最可能规模、最大规模和最小规模。通过这三个值的计算可得到一个统计学上的期望值和一个标准偏差。

公式：PERT 公式估计的预期规模是 E，标准偏差是 SD:

$$E = \frac{a + 4b + c}{6}$$

$$SD = \frac{c - a}{6}$$

举例说明，假如一个新的通信程序：

a=最小规模 例如： 1000 条测试用例

b=最可能规模 例如： 1200 条测试用例

c=最大规模 例如： 1500 条测试用例

那么根据 PERT 公式估计的预期规模是 E，标准偏差是 SD：

$$E = \frac{1000 + 4 * 1200 + 1500}{6} = 1217.6 \text{ 条测试用例}$$

$$SD = \frac{1500 - 1000}{6} = 83.3 \text{ 条测试用例}$$

平均规模 $E = 1217.6$ 条测试用例， $1\sigma = 83.3$ 条测试用例				
范围		概率	从	到
E	$\pm 1\sigma$	68.26%	1134.3	1300.9
E	$\pm 2\sigma$	95.46%	1051.0	1384.2
E	$\pm 3\sigma$	99.73%	967.7	1467.5

这就是说，测试用例规模在 1134.3 和 1300.9 之间的概率是 68.26%，在 1051.0 和 1384.2 之间的概率是 95.46%，在 967.7 和 1467.5 的概率是 99.73%。本估计方法的前提是对规模的估计没有偏见，经验表明，估计偏低的倾向大于偏高的倾向，使用时应加以考虑。

3.3 类比法（Analogy Approach）

当待估计项目与已完成项目在应用、环境和复杂度方面相类似时，可以使用本估计方法。

本估计法的基本步骤如下：

1. 从软件过程数据库和文档库中找到类似项目的相关估计数据和文档；
2. 列出参考项目中可类比的功能和完成这些功能的测试用例；
3. 分析待估计项目和参考项目可类比功能点之间的差异；
4. 考虑待估计项目和参考项目功能复杂度的差异；
5. 依据 2、3、4 步的结果进行估计，形成对项目测试用例规模的估计。

很明显，这种估计的准确性依赖于参考项目的完成程度和数据的准确程度，因此使用这种估计方法要求有一个内容丰富、准确、可靠的软件过程数据库。

3.4 运算法（Algorithmic Approach）

3.4.1 基于代码规模的运算法

本估计方法仅适用于测试团队已经知道被测试项目代码行规模，并且组织 PDB 中有类似项目的测试用例密度，即测试用例总数与项目规模（KLOC）之比。

在进行估计时，有自底向上或自顶向下两种方式进行选择。自底向上是对任务分解结果中不可继续分解的子节点（模块或子模块）进行测试用例规模估计，通过计算每个模块或子模块来生成整个项目的测试用例规模。自顶向下是从整个项目开始，对项目总测试用例规模做出估计，再按照比例或相关关系计算每个模块或子模块的测试用例规模。

计算测试用例规模时，测试用例规模=产品或产品组件代码行规模（KLOC）* 测试用例密度（条/KLOC）。

3.4.2 基于需求抽样的运算法

本估计方法仅适用于测试团队已经知道被测试项目需求规模。

1、对需求进行分类。

对需要开发测试用例的需求进行合适的分类并确定每类需求数量，比如按照功能需求、数据需求（计算功能的相关数据约定）、性能需求（执行速度、精度、可靠性、可维护性、可扩充性、互操作性、可重用性等）进行分类。

2、对分类后的需求进行抽样。

对分类后的需求进行抽样时，需要考虑需求重要性和优先级等属性和项目实际进度要求，按照 5%-10%比例进行样本抽取。

3、对抽样后的需求进行测试用例开发。

分别对抽样后的不同分类需求，挑选比较有经验的测试用例设计人员进行开发，开发完成后，针对不同分类需求，计算每个需求达到测试覆盖要求所需要的平均测试用例数。

4、估计需要新开发测试用例的规模。

依据需求抽样时的比率，使用上一步得出的平均测试用例数，计算需要进行新开发的测

试用例规模，即新开发测试用例数= $A * \sum_{i=1}^n X_i Y_i$ ，($A=1/P$)；其中

A ：表示需求平衡系数，主要是平衡由需求不稳定（需求变更，包括需求增加，需求修改，需求删除）导致对测试用例规模的影响；

P ：表示稳定需求数占总需求数的百分比；

X_i ：表示某一类需求数量；

Y_i : 表示该类需求所需要的平均测试用例数;

i : 表示需求分类的数目;

此外还需要考虑需求抽样, 对估计产生的影响。

4. 附录二：工作量估计的方法

关于工作量估计的方法有很多，下面简单介绍其中的几种方法，但使用时并不限于这几种方法，可以使用其他的方法：

4.1 运算法 (Algorithmic Approach)

运算法是一种简单直观的估计方法，它根据项目测试用例规模估计的结果、历史项目数据，进行运算得到工作量估计：

$$\text{工作量} = P * S$$

S = 测试用例规模（条）

P = 生产力系数（1/每人日测试用例条数）

项目的产品复杂程度、人员能力、前例性、重用比例、考虑未来重用的比例等估计假设条件将影响到项目工作量的大小。根据项目具体情况，设置调整系数。一般地，产品的复杂程度越高，工作量越大；人员能力越弱，工作量越大；以前未做过类似的项目，考虑到一些无法预料因素的影响，则工作量应偏大；如果测试用例重用要求高，为完成此项要求项目的工作量将增加。上述几种估计假设条件，在相反的情况下，工作量应偏少。

4.2 专家判断法 (Expert Judgement Approach)

本估计方法需要使用者有类似系统的经验，结合规模估计的结果和 WBS 综合分析，形成对工作量的估计，可以按照下面的步骤进行：

1. 结合 WBS，根据经验和规模来估计 WBS 中每项原子功能的工作量。
2. 根据每个功能点测试难度和风险的大小，适当调整工作量的估计
3. 按阶段合计需要的工作量，形成总的工作量的估计