

手机测试思想

倪维华

零 写在前面

首先声明，虽然名为思想，但绝没有灌输之意。

总以为自己的见解是正确的，总没有机会把它完整的表达出来，现在就借“思想”的名义，把它表达并记录下来。看完之后，感觉不太认可的地方，请不吝指出，一起讨论以求有个正确的东西大家可以依行。

注：本文所说的测试，若无特别声明，都是指系统测试（即 ST）。

一 应有的观念和心态

手机开发关键的三个要素是：质量、成本和纳期（需要说明一下这三个要素的重要度是依次递减的）。与手机测试直接相关的就有两个要素——质量和纳期。从目前 X300 的情况看，测出的问题没有解决，就将直接影响手机的质量；为了保证产品质量，就得延长开发期限，因而只能推迟上市日期，这也造成了成本的增加。由此可见，手机测试在整个手机开发流程中占据重要的位置，只要测试者还说不行，手机就无法上市（测试者简直是手握生杀大权）。

毫无疑问，要提高手机质量就离不开测试，但测试本身无法直接提高手机质量（用专家的话说就是只能有助于质量的提高）提高质量还要依赖于开发者的设计以及对发现问题的解决。所以，测试的目的就不仅仅是发现问题，而是要有助于解决问题。如果从主席的《矛盾论》出发看问题的话，测试者与开发者对于手机内存在的具体问题，他们是对立的；但是对于提高手机质量的目标他们应该是统一的，只有手机的质量提高了，受到市场的认可，他们对问题的发现与解决才是有意义的。如果测试者发现并提交给开发者的许多问题最终没有被解决，则手机质量就无法保证，就得不到市场的认可，测试者的所做工作也就失去了意义，价值也就无从体现。

好了，说完了大道理，就该说说具体应该怎么做了。我们所从事的系统测试（ST）及现场测试（FT）都属于“黑盒测试”，就是预先设计好一些输入（操作），观察手机的输出（结果或现象），从而判定其是否存在问题。无论做得多么完善的测试式样书，都不可能涵盖所有的输入情况，特别是一些测试过程中出现的随机情况，所以就要考虑利用重要的有代表性的输入情况来测试。这里引用一个等价区间的概念，在等价区间中，只需用任意值测试一次即可。等价区间的概念可表述如下：

记 (A, B) 是命题 $f(x)$ 的一个等价区间，在 (A, B) 中任意取 x_1 进行测试。

如果 $f(x_1)$ 错误，那么 $f(x)$ 在整个 (A, B) 区间都将出错。

如果 $f(x_1)$ 正确，那么 $f(x)$ 在整个 (A, B) 区间都将正确。

上述测试方法称为等价测试，来源于人们的直觉与经验，可令测试事半功倍。一般来说，在作成测试式样书时，各种等价区间及分界值都已经被考虑在其中了，这里不是要讨论如何设定等价区间，而是要说如何利用这种思考方式去再现问题。举个例子，有一次在做 MMS 测试时，发现在播放带有 gif 格式动画的 MMS 消息时，手机自动重启了（自动重启是个很严重的问题），如果到此为止就向开发者报告这个问题的话，这时的等价区间是所有带有 gif 动画的 MMS，当开发者要解决这个问题时就会变得很头大，因为他们发现有时带有 gif 动画的 MMS 也不会重启，那到底问题在哪呢？只好发回要求重测。如果测试者在发现问题时，能够继续挖掘更深层的事发原因、大大缩小等价区间的话，开发者就能发现问题所在并能很快解决问题。还好，刚才的问题当初是这么处理的：发现问题后，充分发挥想象力和耐心，继续试验各种大小的 gif 文件，设置不同的页面时间，最后发现当页面时间大于 20 秒时，播放带有 gif 动画的 MMS 一定

会导致机器重启。此时的等价区间可以说是大大减小了，结果问题很快被解决。

这里还想插入一些闲话，以前有人发现问题（等价区间很大）后不假思索就要发报告，我让他够继续挖掘一下更深层的原因，他告诉我他又不知道源代码怎么知道深层原因（我倒！），现在有了等价区间的概念，希望不再出这样的笑话。还有，有人认为是为了挖掘更深层的事发原因、缩小等价区间，要花很多时间，影响测试进度。对于这一点，我是这么认为的：对于测试结果本身它应该也是有质量要求的，同样一个问题，如果报告时等价区间很大，开发者不易解决，显然质量是低的；相反，如果等价区间很小，开发者很容易解决，他的质量是高的。质量要素应该优先于时间要素，就像为了提高质量而推迟手机的上市日期一样。另外，如果想在短时间内缩短等价区间的话，就要利用自身的经验并且充分发挥想象力，然后通过操作去验证。总之，我们的测试及问题的报告要有助于解决问题

再说“技术者良心”的问题。做人要讲良心，就是要明是非、判正误、晓情理；作一个技术者，就要讲“技术者良心”，就是要对自己所做的技术活讲良心。具体到测试者，就是要对测试的结果和发现的问题讲良心，在测试中遇到的任何异常情况，都应该积极去探究其是否是一个问题，不因为它难再现、难描述而放过（这样往往会把问题留在手机中让用户去发现）；对于那些明显没问题的测试项目，也要不遗余力地去确认一遍。每当我想偷懒的时候，我将感到深深的惶恐，万一这个地方出问题怎么办？

通常手机软件 BUG 的再现及解决状况呈如图 1-1 所示的曲线分布，每一个型号的手机根据其开发量可以估计出其 BUG 总量，通过测试必须再现出九成以上的 BUG 后才能结束测试；开发者解决了所有发现的 BUG 并经过发布测试（即 RT）之后才能正式发布上市。所以，BUG 发现得越晚，对应的解决也越晚，如果致命的问题总是很迟才发现，那必然就要影响手机的纳期了。所以，有问题应该尽早地报告，而不是积压下来择日再发。

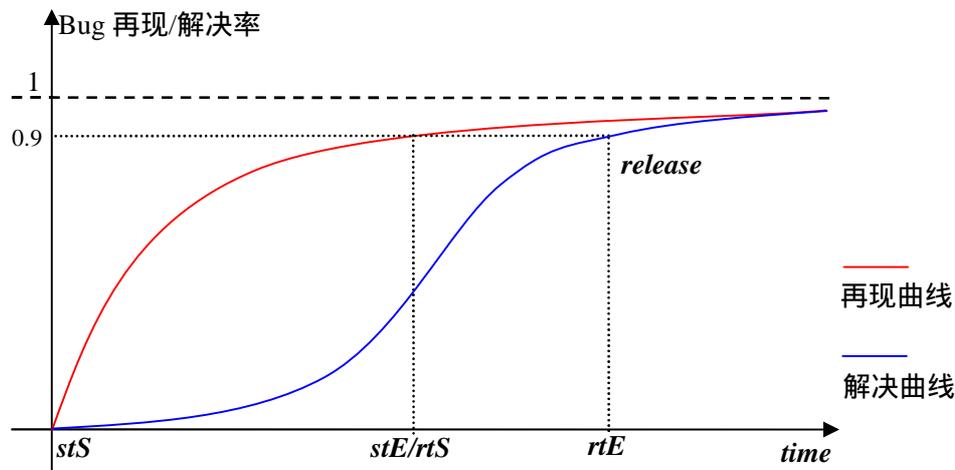


图 1-1 Bug 再现/解决状况曲线图

最后一点，就是手机软件的 Bug 没有完全被发现并解决的那一天，只要有时间继续测试，总可以发现手机里存在的问题，无论它是质量多么好的手机（从“只要是程序，它就有 Bug”得出的推论）。因此，争取在手机发布之前把所有致命的问题（重要度再 A 级以上）都再现出来，并解决掉，是我们测试的直接目的。对待致命的问题要像对待“非典”一样，“早发现、早隔离、早诊断、早治疗”，这样才能使手机质量有保障，才不会影响纳期。所以测试过程就要讲究一个轻重缓急，先后有序，轻的缓的测试让初级的测试员（后面会具体谈到测试员的分级）来做，重的急的就得交给高级的测试员。

二 测试员的等级

日语考试要分级，程序员考试要分级，所以测试员也要分级。这样才有利于测试任务的分配及工作流程的管理，才有利于测试员自身素质的提高，才有利于整个测试行业的良性发展。为此，针对于手机的测试，将测试员分成以下几个级别。

初级测试员 : (reproduce routine bug)

1. 熟悉手机操作；
2. 了解当型号手机的总体功能；
3. 能读懂测试式样书并根据其进行操作；
4. 能够报告与式样测试点不符的问题。

中级测试员 : (reproduce random bug)

1. 具备初级测试员的能力；
2. 了解 GSM、gprs 及其他相关的协议；
3. 能够利用各种工具（烧录工具、取 Log 工具、超级终端等）；
4. 能够发现并报告测试式样以外的问题及取 Log。

高级测试员 : (confirm bug)

1. 具备中级测试员的能力；
2. 对于 Phone Function、MMI、Message、Wap、Java、DRM、File System、Camera、Gprs 等机能，掌握某一方面（或几方面）的原理；
3. 能够编写或 Review 精通方面的测试式样书；
4. 于精通方面能够确认问题是否存在。

某型号测试主担当 (manage bug)

1. 是某一方面的高级测试员；
2. 能够根据测试员的等级分配所有的测试任务，保证测试进度；
3. 了解并跟踪当型号出现的所有问题；
4. 能对当型号手机质量给出整体评价。

大家不要有误解，这里给测试员分等级就是为了给测试员一个名分，实际上这与名分基本上没有什么关系，而主要是从能够承担的责任及分担的工作上来划分的，目的就是为了更快更好的完成测试任务。同时，它也为测试员的努力方向提供目标参考，使针对测试员的培训更加有的放矢。在具体操作中，对于新加入测试员行列的人来说，它只能从初级测试员做起（不论他在从事其他方面时是多么优秀），在完成一个型号的初级测试任务后，就可以考虑让其承担中级测试任务，中级测试员在完成了一两个型号的中级测试任务后，就可以考虑分担高级测试任务，最后，对于有耐心和细致的高级测试员，才能委以型号测试主担当的工作。当然，还得看各自自身的努力情况而定。

三 测试的内容及分工

手机测试工作，说它简单，它的却可以是一个比较简单的工作，如果测试式样书科学而且足够细致、测试条件充分满足，只需要测试员一步步地操作并确认最终结果而不需要关心其他的事情；但是说它难，它也会是一个很难的工作，如果测试式样书不是那么的完美，而且还有许多过程中的异常情况需要捕捉，更要命的是碰到一些情况无法判定它是对还是错，比

如说正在进行 gprs 数据传输时，无法打进电话)，这时测试将变得很难，需要很多的调查和比较。

手机测试的内容一般会按机能来划分，每一项机能独立进行测试，比如前面提到的手机的 9 大机能，当然并不是所有型号的手机都具备这些机能，测试内容也可以随之增减。对于某一项特定机能，又该如何测试及分工呢？下面以 WAP 机能为例，具体叙述。

WAP (Wireless Application Protocol) 是无限应用协议的简称，主要是为了给无限移动终端上提供一种网络浏览的应用，相当于提供给计算机终端的 WWW 应用。协议里详细地规定了这种应用是如何实现的，包括：WAP 站点是如何建立的、如何通过 WAP 网关接入到互联网，WAP 浏览器又具备怎样的功能，承载服务的数据是按什么格式传输的等等。WAP 应用具体到手机上的 WAP 机能就是 WAP 浏览器，它相当于计算机上的网页浏览器 (IE 或 Netscape)。测试这个 WAP 浏览器，主要从两方面即表现系和操作系来考察。表现系就是要测试浏览器能否正确表现出各种语言 (WML、XHTML、CHTML 等) 编写的 WAP 网页，这部分测试属于比较容易的操作，可以由初级测试员来完成，前提是各种语言的 WAP 网页准备齐全，式样书中明确记述各个确认点。另一个操作系是要测试浏览器的各种操作，包括：连接设置、浏览器设置、菜单操作、PUSH 设置及操作、以及在进行 WAP 操作时与优先级高的事件的竞争测试。这部分的测试就相对复杂，除了要确认测试点外，还要留意操作过程中的一切异常情况，因此，测试就该有中级测试员来完成，部分操作如 PUSH 测试、竞争测试还需要由高级测试员来完成。这里的分工主要从提高测试效率和测试效果两方面来考虑。

每项手机机能的测试总是可以划分出简单的测试和复杂的测试，分配给不同等级的测试员来完成。如果在分担任务时，只考虑测试项目的数量而不考虑项目的难易，将不利于提高测试效率和效果。

四 问题报告

“怎样的报告才算是好的高质量的报告？”我总在问自己，具体详细的？描述清晰的？这些都是必要条件而不是充分条件，我觉得最有助于开发者解决问题的报告才是最好的报告。当然，好的报告离不开认真细致的测试过程。

先说不好的报告会带来什么结果。首先是开发者看不懂，不理解问题出在哪，也就无从下手解决；然后他会发回要求重测或需要更多信息；接着测试者就再来一遍；若提交的报告还是没有多大改进，开发者继续要求取 Log 或重做多少遍，来来往往浪费时间不说，测试者也会被弄得疲于奔命、不堪重负。

现在 ODM 手机开发方式采取的是开发方与测试方分立而行，这种做法对于提高手机质量总体上是利大于弊，但可能造成一种不好的趋势，就是使开发者变得越来越懒惰，写完程序，不愿做太多的走码测试就交给测试方去做测试，这样出问题的机会就很多，出的问题也很怪。因此，测试者就更要讲“技术者良心”，测试时认真细致，提交报告时多一点聪明，让开发者尽快地解决问题，测试者自己的日子也会好过一些。

为了提高报告的质量，报告提交前进行确认是必要的，由高级测试员来确认中级、初级测试员的报告，由测试主担当来检查并跟踪所有的报告，这样就能做到分工明确、责任清晰。总之一句话，就是既要快速积极的报告问题，又要不使低质量的报告流放出去、遗害人间。

好了，以上就是想表达的东西。