移动应用自动化测试文献综述

姓名：丁晓勤 学号：171250005

1. 引言：

近年来，随着移动设备的不断普及，基于移动设备的移动应用程序在覆盖范围、种类和数量上都在急剧增长，移动应用程序涉及的领域也从小而无用的娱乐工具逐步向流行且常用的媒体社交工具过渡，甚至了在更具专业性且有更高要求的商业、生物医学方面也有所涉猎，这些都使得移动应用测试变得尤为重要。由于移动设备的规格、系统种类的多样性以及测试的日益重要，传统的测试方法已经无法满足移动应用的正确性和快速发展，这要求我们转而寻求更快捷、更省力的测试方法。从收集阅读的文献中可以找到许多种测试方法，但是尚不能确定的是这其中提出的测试方法是否是符合工业或者某一领域的移动应用的，是否能在测试的有效性和效率方面达到符合标准的效果。为了找到这些问题的答案，本次文献综述打算从以下几个方面出发，探索移动应用与测试方法的相关性以及找出更符合标准且更切实可行的相应的测试方法。确切的研究问题如下:

·RQ1：与常规应用程序相比，移动应用程序有哪些不同？相应的移动应用测试存在什么样的待解决问题？

·RQ2：传统应用和移动应用的测试方法的比较？

·RQ3：移动应用测试未来的方向是怎样的？未来还将出现哪些问题？

下文第二部分（背景）中会对问题RQ1给出回答，在第二部分（基本内容）中会对RQ2中提出的问题进行详细分析。

1. 正文：
   1. 研究方法分析：

在本节中，描述如何进行文献调查。

通过阅读作业要求，选取可以使用的关键词和搜索条件，在ACM数字图书馆、SpringerLink、IEEE中搜索关键词，阅读初始论文并决定研究方向，根据研究方向关键词进行二次搜索。主要选取期刊文章、有关移动应用自动化测试和众包测试论文等。

参考文献可以在第五部分附录中查看。

* 1. 背景：

从1994年IBM Simon的发布，再到2007年第一款iPhone的发行，智能手机才渐渐走进我们的视线。装载移动应用的移动设备渐渐从智能手机衍生到了平板电脑、智能手表，移动应用也从最初的增强型娱乐应用发展到后来承载了音、视、乐、生各式业务的多样化应用。相较于传统应用，移动应用以其更多的数量、更快的速度、更便捷的使用、更多样的领域占据我们生活的方方面面。移动应用已经嵌入了我们的生活，这使得我们对移动应用的有效性和稳定性提出了更高的要求。

与有效性和稳定性要求相悖的，是移动设备的规格多样、各厂商的操作系统的差别之多，这也是移动应用程序与常规应用程序最大的不同之处。

三种主要的移动操作系统：Android，iOS，Windows Phone。根据市场净份额，Android是2016年第一季度最受欢迎的操作系统，市场份额为60%，其次是占比32%的iOS，Windows Phone以3%排名第三。Android系统由于它的免费开源受到广大移动设备制造商的欢迎，这也是Android系统市场份额如此之大的原因所在，但是正由于它的免费开源，各厂商都热衷于对其做出细微修改之后再装载在相应机型上。移动设备相较于传统的台式电脑来说，需求量更大、代价更低、更新迭代更快，这不仅使操作系统的变化更加频繁，也使移动应用的更新加速。与常规应用程序相比，移动应用程序最大的不同之处在于它的碎片化严重，不论是同等Android操作系统之间的差别，也是尺寸、规格、物理硬件上的碎片化。

移动应用的数量之大、碎片化严重都对测试方法和人力提出了更高的要求，为了寻求效率和有效性的折衷，我们需要对传统测试方法和文献提出的新的测试方法做类比，寻求更好的解决方法。

* 1. 基本内容：

传统的程序测试根据项目流程阶段划分测试，典型的单元测试、集成测试、系统测试、验收测试、回归测试。传统的程序测试强调验证程序的功能性，且相较于移动应用而言，程序的独立性更强，应用与应用之间的依赖性较低，这使程序无需考虑各种应用之间的接口调用。另一方面来说，传统应用程序依赖的硬件设备更为稳定，由于传统应用程序的场景以及功能的固定性，其需要测试的场景更少，测试的需求功能也更少。一般而言，测试开发人员更新维护周期更长。

相较于传统的程序测试，移动应用测试除了需要考虑基本的功能需求测试，还有许多其他问题需要考虑。

首先，上文提到的移动设备的多样性以及应用所依赖的操作系统之间的差别都对移动应用程序的兼容性提出了巨大的挑战。

其次，要考虑其他应用程序对本应用程序的影响，尤其是移动设备基础的社交通信应用的影响。其他应用的影响所产生的问题可能是致命的，尤其是导致中断、功能缺失这样的问题时。

另一方面，界面的美观和舒适、人机交互的便捷是一个移动应用长久存活的重要因素，移动设备规模尺寸的不同对UI的影响也是巨大的，这也是测试的一个难点。

由于移动应用的数量之多，测试人员很难在首次开发测试阶段排除所有应用产生的影响，这就要求测试人员在应用向用户开放后不断完善，不仅对测试人员的数量提出了要求，对用户配合度也提出了新的要求。

已经提出了许多关于移动应用测试的不同框架和工具。Baride和Dutta用基于云的方法来解决不同平台上的移动应用程序的差异问题，Ping等人提出了基于V模型和云测试机制的框架，为移动应用程序测试设计模型提供了标准，包括使用网络连接、测试范围等。

移动应用测试和传统应用测试的前期阶段都是相同的，都是基于基础测试的，关于设计测试用例的方法（边界值分析、等价类划分等）、测试方法（黑盒测试、白盒测试）、检查UI的标准（布局、风格、按钮等）。相较于传统应用测试，移动应用对测试的范围提出了更高的要求，要满足各种中断以及新旧版本的覆盖要求。移动应用自动化测试工具、移动应用众包测试也应运而生。

自动化测试是测试人员使用测试工具运行测试脚本的过程。通过减少测试周期以及提高生产率和质量，使用工具来帮助测试过程，包括：选择测试工具，定义自动化，计划，设计和开发，测试执行和维护的范围。

OpenSignal的最新研究中显示2015年8月超过24,000种支持Android [1的](https://link.springer.com/article/10.1007%2Fs11219-018-9418-6#Fn1)不同类型的设备已经存在。就效率和有效性而言，随着移动设备和移动应用的不断发展、移动设备的碎片化问题，传统的手动测试移动应用程序既昂贵又枯燥、且不能保证正确性。测试自动化正在变成提高移动应用程序质量并降低测试成本的有效解决方案。就Android而言，已经发布了Monkey、MonkeyRunner、Appium。

对于移动应用的自动化测试，在便捷和效率的优点之后，也存在很多等待解决的问题。一方面，移动应用自动化测试工具的开发相较于传统移动应用测试而言是极其困难的，对开发人员能力的要求更高，不仅在于软件编程方面的能力要求，更可能涉及商业、医学等专业知识的要求。另一方面，对自动化测试工具的正确性验证也是一个极大的考验。严格来讲，某些问题更容易被错误地回答，工具提取的数据和纳入和排除的标准的正确性无法保证，分类方案的适用性、结论的有效性和测试的安全性都很难被满足。

众包移动应用测试通过更多的“人”来进行测试，与传统的移动应用测试的不同在于，众包测试对测试人员能力的要求较低，对人的数量和范围的要求更大更广，是属于一种线上各自独立的测试方法，众包测试人员无需和其他众包测试人员进行交流。众包测试一部分弥补了自动化测试专业性不强的缺陷，相比来说，对人机交互部分、异构设备和场景的测试能力更强，因为是通过真实且大基数的人来进行测试的。同样地，众包移动应用测试也有其自身的缺陷，首先，数据的大规模共享使得我们无法保证其安全性。更大的问题在于，由普通用户提供的测试报告的质量和可靠性无法保证，测试数据可能具有高度重复性，大基数且不专业的测试报告会导致离散的数据，这对测试数据的筛选和处理提出了更高的要求。

1. 结论：

传统的应用测试方法已经不适用于不断更新迭代的移动应用，在选取移动应用测试的方法时不能一概而论。对于小而简单且娱乐性较强的移动应用，传统的测试方法足以满足测试要求。对于大规模准确性较强的，应该转而选用自动化测试工具，对于专业性较强的应用测试，传统测试方法与众包测试的结合更为适合。随着移动设备涉及的感官领域的增加，单一的移动应用测试方法将不能满足测试要求，尤其是音、视、声、触等的加入与混合，这要求测试人员用更新、更多的方法来解决所遇到的测试问题，例如自动化测试与众包测试的结合测试。

1. 附录：

[1]Porfirio Tramontana,Domenico Amalfitano,Nicola Amatucci,etc.Automated functional testing of mobile applications: a systematic mapping study[J].Software Quality Journal,2019,27(1):149-201.

[2]Shkodran Zogaj,Ulrich Bretschneider,Jan Marco Leimeister.Managing crowdsourced software testing: a case study based insight on the challenges of a crowdsourcing intermediary[J].Journal of Business Economics,2014,84(3):375-405.

[3]Ruizhi Gao,Yabin Wang,Yang Feng,etc.Successes, challenges, and rethinking – an industrial investigation on crowdsourced mobile application testing[J].Empirical Software Engineering,2019,24(2):537-561.

[4]Diego F. Aranha,Helder Ribeiro,André Luis Ogando Paraense.Crowdsourced integrity verification of election results[J].Annals of Telecommunications,2016,71(7-8):287-297.

[5]Luis Cruz,Rui Abreu,David Lo.To the attention of mobile software developers: guess what, test your app![J].Empirical Software Engineering,2019,24(4):2438-2468.

[6]Chien-Hung Liu.A compatibility testing platform for android multimedia applications[J].Multimedia Tools and Applications,2019,78(4):4885-4905.

[7]Mamta PandeyRatnesh Litoriya,Prateek Pandey.Novel Approach for Mobile Based App Development Incorporating MAAF[J].Wireless Personal Communications,2019,107(4):1687-1708.

[8]Sergiy Vilkomir.Multi-device coverage testing of mobile applications[J].Software Quality Journal,2018,26(2):197-215.

[9]Triin Samuel，Dietmar Pfahl.Problems and Solutions in Mobile Application Testing[EB/OL].International Conference on Product-Focused Software Process Improvement,2016-11-6.

[10]Maryam Ahmed，Rosziati Ibrahim.A Comparative Study of Web Application Testing and Mobile Application Testing[EB/OL].Advanced Computer and Communication Engineering Technology,2014-11-2.

[11]Hao Lian，Zemin Qin，Hangcheng Song，Tieke He.E-CAT: Evaluating Crowdsourced Android Testing[EB/OL].International Conference of Pioneering Computer Scientists, Engineers and Educators,2018-9-9.

[12]Bin Guo，Huihui Chen，Yan Liu，Chao Chen，Qi Han，Zhiwen Yu.From crowdsourcing to crowdmining: using implicit human intelligence for better understanding of crowdsourced data[EB/OL]. <https://link.springer.com/journal/11280>, 2019-8-31.