Application of software testing methodology based on quality criteria and expert assessments to mobile applications

E Y Galimova
Saint-Petersburg State University of Industrial Technologies and Design, Saint-Petersburg, Russian Federation
E-mail: galim81@mail.ru

Abstract. In this paper, we propose a technique for choosing between a manual, automated and semi-automated approaches to software testing. The great desire of each person to achieve maximum comfort in each of the spheres of life has also affected the international web. The user, wanting to always stay online, uses the phone as a communicator. This led to the emergence of mobile Internet. When staying away from home or during travel and business trips, you can connect using a tablet or mobile phone instead of a laptop. The efficiency and functionality of “mini” computers would not be brought to such a high level without specialized applications. Properly chosen testing methodology improves the quality of mobile applications.

1. Introduction
In the modern world, mobile devices are equipped with almost the same set of functions as stationary computers [1]. A distinctive feature of mobile devices is a more complex interaction with users. This fact makes their testing more resource-intensive [2], and organization should efficiently plan the testing process.

Quality assurance issues have been discussed in a number of research papers [3, 4]. The quality of data in software systems must comply with the characteristics described in the standards of the ISO 8000 group. Software testing is part of the quality assurance process for software systems. The quality management system for software systems should include all stages of the life cycle [5, 6]. This article proposes a formalization of the overall assessment of a software system for testing purposes. This is done by answering a list of questions that are close to unambiguousness to solve the issues of organizing testing of a given software system [7]. Experts are invited to fill out a questionnaire. The questions are based on the IEEE 830 standard, which forms the requirements for the software system. Examples of questions from the questionnaire are presented below.

2. Automated testing criteria

Do you plan to use cloud services for testing? Cloud-based testing platforms can be scaled up to many operation systems and devices. Mobile applications should work correctly on various combinations of platforms [8]. Cloud test automation tools are designed for automated test data generation, test case generation, test execution and test evaluation [9]. Consider the testing algorithm using a cloud farm with DevOps [10]. The last branch from the repository is loaded into the build system and the application is compiled. During the compilation process, you can run unit tests. Next the assembly is uploaded to the cloud device farm, where automated user interface testing (User Interface, UI) will be performed.

Is random testing important? Testing done without any plan; tests are run in random order. For example, clicking in a chaotic order on the first available buttons in the application.
applications for automation. For Android - UI / Application Exerciser Monkey, a command line tool that can be run on any emulator instance or device. It sends a pseudo-random stream of custom events to the system. The application has a number of settings: basic configuration parameters, operational limits, event types and their frequency, debugger settings.

**Is penetration testing required?** The purpose of this type of testing is to try to obtain confidential information; to use confidential information to commit malicious acts; to become a “superuser”, get bonuses and privileges. Software protections must comply with the requirements of the Mobile Application Security Verification Standard (MASVS). There is a research open source community dealing with issues of application vulnerabilities (iMAS). Created automated tools for penetration testing of mobile environments (MobiSec, OWASP Zed Attack Proxy Project, IBM Application Security on Cloud) [11].

**Is smoke testing planned?** Post-build software testing to ensure critical functionality is working properly. Performed "before" any detailed functional or regression tests [12].

**Will data quality testing be performed?** The testing process verifies that incorrect data is processed correctly.

**Do you plan to search for challenges of Transport Layer Security (TLS)?** Starting with the Android PieE version, all incoming and outgoing connections will be made via the TLS protocol. It prevents unauthorized access and listening to packets when exchanging data in client-server applications. At the same time TLS does not allow lowering the version of the security protocol and weakening the encryption algorithm. Open source tool Nogotofail designed for testing TLS and works for Android, Windows, Linux, iOS, OSX, Chrome OS. The UniTEST automated testing technology can be used to verify TLS security protocol verification [13].

3. **Manual testing criteria**

**Do you plan to use alpha testing?** This is manual testing, one of the most common approaches to testing software systems. Used in the development of software systems, towards the end of the development stage. Consists of several iterations. Performed by the developers themselves, internal testing specialists. The advantage of alpha testing is that the found defects can be quickly eliminated by the developers. The testing process is carried out by performing tasks that a potential user can perform. Alpha testing can act as internal acceptance testing for the finished software system prior to beta testing. It is carried out using the white, gray or black box method.

**Do you plan to use beta testing?** This manual testing is done to get feedback from real users. Usually done by the black box method. Beta testing does not require a dedicated lab or test environment to be deployed. Testing is carried out in a real time environment. There are closed and open beta versions. The closed one is issued for a limited number of users, the open one is available to everyone interested.

**Is installation testing important?** Check that all previous data is saved after updates. Updates should be installed via GSM-connection, Bluetooth connection, via Wi-Fi. Possible installation options for mobile application shown in Figure 1.

![Figure 1. Types of mobile application installation testing.](image)

**Will the touch interface be tested?** You need to check that all areas of the screen are equally sensitive to pressure. Check, how much effort you need to make to control the screen and whether light touches are registered.
Is accessibility testing important? The application should work correctly with at least some of the following functions that increase accessibility: High Contrast Display Modes, Voice Recognition, Screen Keyboards, Peripheral Keyboards, Blinking cursor, Peripheral Mouse, Visual interpretation of sound alerts.

Is interrupt testing planned? Different types of mobile interruptions shown in Figure 2.

Figure 2. Basic test cases for mobile interrupt testing.

It is important to distinguish between third-party application and system notifications, since they are processed differently. Third-party applications allow you to quickly emulate the notification; system applications can notify about insufficient charge or availability of updates, and their notification is a bit more difficult to emulate. Tester must to check the transition to the standby mode after the user intentionally locks the screen. Like all checks, this should be done while working with the main functionality of the application, since locking can significantly affect the processing of processes.

Is economic gain important in the short term? The cost of manual testing depends only on the salary of the tester. The company does not need to pay for automation tools. Manual testing is suitable primarily for small projects. A number of approaches have been developed for calculating the efficiency of automation applications [14].

4. Semi-automated testing criteria

Will we test the accelerometer? With the development of technology, almost all mobile devices are equipped with an accelerometer. The main purpose of this sensor is to provide information about the current acceleration of the device, more precisely, the difference between the acceleration of the device and the acceleration of gravity. At rest, the sensor readings coincide with the gravitational acceleration vector. The motion recognition algorithm using an accelerometer is described in [15].

Owing to the accelerometer, the mobile device responds to changes in orientation in space (turning the display, shaking, hitting), which will allow the screen to be rotated at the right time, as it would be convenient for the user. The accelerometer is used when operating the device as a pedometer. Calibration of the accelerometer can be done manually using the GPS Status & Toolbox for Android. You can record and play back the sequence of events from the accelerometer using the Android Debug Bridge (ADB) using the GetEvent / SendEvent commands.

At the device level, accelerometer testing is usually done using a microcontroller unit (MCU). Accelerometer and MCU are soldered to a printed circuit board. To verify the accuracy of the acceleration data received from the accelerometer, a special test program is developed for the MCU. Accelerations are read from each of the three axes. The obtained values are transferred to a computer for analysis.

Is regression testing planned? During testing, the existing functional and non-functional areas are checked after making changes to a part of the software or after adding new functions.

Is formal testing important? Standardized verification of software products, performed in accordance with the test plan and other formal documentation. During the testing process, final tests are performed for all the functionality of the software system. Formal mathematical methods are applied.
**Is testing of responsive design planned?** Adaptive design provides good perception when viewing pages. The site should adapt to the user mobile device. Adaptive design provides good perception when viewing pages. The site should adapt to the user gadget. Manual checks are carried out on existing mobile devices. Next, we use tools that mimic the screens of various mobile devices. For example, Am I Responsive allows you to drag and drop devices anywhere on the screen, you can embed your test site in a link and share it. Only four types of screens are available: desktop monitor, laptop, tablet, mobile phone. In another application, there are more device options and screen resolutions. There is a function to edit the background header. You can use the mobile emulator from Google DevTools. Without the need to install adb-drivers or something to configure, desktop Chrome detects all running instances of Chrome and WebView in Android and displays them on the screen, providing the ability to manage them both from the phone itself and from the desktop browser.

**Is the test object a hardware and software complex?** Usually, the main components are tested separately, then manually analyze the results obtained. Next is a semi-automated comprehensive testing of the entire system. Consider the software and hardware complex on the example of digital technological radio systems. Single special standard is used for rail transport. It is called the global mobile communications system for rail transport (GSM-R). There is also a backup radio standard Digital Mobile Radio (DMR) [16]. Scheme of digital technology radio system for GSM-R shown in Figure 3.

![Figure 3. Digital technology radio system.](image)

Mobile station (MS) is a mobile phone or other subscriber equipment that receives GSM communication services. MS consists of a mobile terminal and a SIM card. The mobile terminal has a unique identifier International mobile equipment identity (IMEI). The SIM card has a unique identifier International mobile subscriber identity (IMSI).

Base Transceiver Station (BTS) provides radio communication for a specific area. BTS installation options: on special masts, on the roofs of buildings, on boiler tubes. BTS functions: provision of interface with Base Station Controller (BSC); radio link management between the Mobile Station (MS) and the Base Station System (BSS). BTS is a distributed system. For testing of distributed systems we apply unit testing, system testing, integration testing. An additional approach is fault injection. In the process of the system using special programs, failures are added: network failures, failures inside
components, disk failures. BTS must be resistant to these faults. Another complementary approach is formal verification. The description of internal algorithms and protocols is checked for correctness.

The main element of the base station system is the Base Station Controller (BSC). Functions: transport channel control with BTS; transport channel control with Transcoder (TC) or with Transcoding Rate and Adaptation Unit (TRAU); transport channel control between TC and Mobile Switching Center (MSC). We can use the BSC simulator for testing.

Handover - the procedure for transferring the active connection between cells. It allows subscribers to move within the network of the operator without breaking the connection. Many disconnections in GSM are due to Handovers, so very high requirements are placed on radio coverage in GSM-R. We can apply automated testing using handover selector role and handover client role. Client can open, close, reopen a connection with the server correctly.

TC (TRAU) performs the function of conversion of the speech signal from the main digital channel to the stream with a lower speed. For the choice of speed is responsible BSC. We can execute performance testing.

Home Location Register (HLR) stores the part of the location information of a mobile station that allows the central switch to deliver the call [17]. The HLR and MSC provide call routing and location changes (roaming) to the mobile station and contain all the administrative information of each subscriber registered in the corresponding GSM-R network, along with the current location of the mobile stations. We can apply automated load testing.

The Visitor Location Register (VLR) is the second main device that controls the movement of mobile stations from zone to zone. With its help, the operation of mobile stations outside the zone controlled by the register is achieved. When the mobile station moves from the zone of one BSC controller to the range of another, in the process of moving it is registered, and new information is entered into VLR. For network testing we can use Visitor Location Register Emulator.

Equipment Identity Register (EIR) is a database with information about subscriber equipment, indicating whether this equipment can be registered on the network or not. EIR stores three lists (white, gray and black) with IMEI. The presence of IMEI in the white list allows access to the network unconditionally. Equipment from the gray list will be allowed into the network, but will be continuously monitored while it is online. The black list is intended to store IMEI devices that are denied access to the network. In India, e.g., the cellular operators with the help of EIR can search out all mobile phones latched to their network [18]. Standard tools for testing databases are usually supplied by manufacturers. There are open source tools: Database Benchmark, Database Rider, DbFit, DbUnit, and more. If we are testing the white box, we need to test the database schema, including checking each object in the scheme; trigger testing. We can conduct stress testing. When testing the black box, we perform testing through the interface.

The Authentication Center (AUC) consists of several blocks and generates keys and authentication algorithms. With its help, the subscriber’s authority is checked and his access to the network is provided. AUC makes decisions about the parameters of the authentication process and determines the encryption keys based on the database in the Equipment Identity Register (EIR). We can execute security testing.

Operations and Maintenance Center (OMC) provides management of network elements and the quality of its work. From OMS we execute control testing of the entire system as a whole. Triotrace Test Monitor can be used to monitor data and evaluate quality of service information for GSM-R. The Triotrace Test Monitor has Forcing functions for voice codec, channel codec, band codec. You can apply manual stress testing for failure of feeders (auxiliary lines between transformers and certain switches). The first cable, e.g., is considered the main one. We check that when the power is turned off, the magnetic starter commutes the electrical circuit to the backup feeder. In case of failure of two feeders at the same time, a mobile diesel generator should be connected.

5. Selection the software testing method using the multi-criteria task
The problem of choosing a testing method (automated, manual, mixed) is posed and solved as a formal multicriteria problem. In its formulation, this problem is nonclassical; it determines not the best element on a given set, but the best solution (alternative) for a given software system. At the same time, ideology and mathematical methods are used, similar to the solution of multicriteria problems.
To rank the questions in order of importance for making a decision on test automation, experts are invited to fill out tables where the degree of pairwise dominance of questions is assessed according to the scale developed by Thomas Saaty, which includes estimates from 1 to 9 with a single step [19, 20].

Using pairwise comparison of questions, it is possible to compose the most complete reflection of the subjective preferences of experts. With this approach, fewer restrictions are imposed on the choice compared to a number of other types of expert assessment. At each step, the expert makes a choice of only two alternatives. This task has an uncertainty level of 1 bit. The described approach facilitates the work of experts. The number of compared pairs for each questionnaire is \( n \times (n - 1) / 2 \), where \( n \) is the number of questions. Questionnaires with tables of pairwise comparisons of questions from each expert are individually processed according to a special algorithm described in the work of V. V. Davnis [21]. Each questionnaire is a square \( n \times n \) matrix \( A \). The calculation is performed according to the iterative algorithm.

When selecting experts, it is not always possible to create a homogeneous group, therefore it is proposed to calculate the weighting coefficients of the competence of each expert. Let us apply an iterative algorithm that calculates in parallel the group preferences and competence coefficients of experts, considered in the work.

Let us construct a rectangular matrix \( P \) of size \( n \times m \), where \( n \) is the number of questions, \( m \) is the number of experts. Let’s introduce additional designations: \( p_{ij} \) is the \( i \)-th row of the matrix \( P \); \( p_{j} \) is the \( j \)-th column of the matrix \( P \); \( p \left( p_1, p_2, ..., p_n \right) \) is the group evaluation vector; \( v^0 = (v_1^0, v_2^0, ..., v_m^0) \) is a vector of initial competency weights.

In the first step of the algorithm:

\[
y^0 = \left( \frac{1}{m}, \frac{1}{m}, ..., \frac{1}{m} \right).
\]

Next, we calculate:

\[
p^i = v_1^0 p_1 + v_2^0 p_2 + ... + v_m^0 p_m = P v^0.
\]

We specify the weighting factors of experts’ competence:

\[
v^i = p_1 v_1 + p_2 v_2 + ... + p_m v_m.
\]

Next, we carry out the normalization and repeat the calculations using the iterative algorithm of parallel computations to the declared accuracy.

At the next step of the algorithm, a rectangular matrix \( Q \) of size \( n \times m \) is constructed, where \( n \) is the number of questions, \( m \) is the number of experts, into which we write the experts’ answers to 70 questions. The answer “yes” is denoted as 1, the answer “no” as 0. Next, calculate \( q_i \), where \( i \) varies in the range from 1 to \( n \):

\[
q_i = q_{1i} v_1 + q_{2i} v_2 + ... + q_{mi} v_m.
\]

The We get the resulting vector \( q \left( q_1, q_2, ..., q_n \right) \) – the weights of answers to questions, taking into account the coefficients of the experts’ competence. Next, the convolution of the vector \( q \) and the vector of group assessment of questions \( p \left( p_1, p_2, ..., p_n \right) \) is calculated. Depending on the result obtained, a recommendation is made to choose an automated, manual or mixed type of testing [22].

6. Conclusion

The development and testing of mobile applications have become one of the most popular tasks in the field of information technology due to the fact that the availability of mobile devices has made them an indispensable tool for solving a large number of tasks: data exchange, quick access to relevant information, mobile payment services, monitoring the flow of business processes, analytical reports and more. The method described in the article allows the testing engineer to make a decision on choosing an approach to testing a software product. In this paper, we approve the application of the selection method of testing type to mobile software product. Testing mobile applications is a separate direction in ensuring the quality of a software product. An engineer involved in a mobile product testing project should be aware of the approaches to the process, as well as special tools and specific checks. From the foregoing, the universality of the method proposed in the article follows, since it applies to mobile, desktop, and web applications. It also has a software implementation.

References

[1] Allen S, Graupera V and Lundrigan L 2010 Pro Smartphone Cross-Platform Development 1–14
[2] Choi W 2013 *Proc. of the 2013 ACM Workshop on Mobile Development Lifecycle* ser *MobileDeLi* 27–28
[3] Crosby P B 1978 *Quality is free* (US: McGraw-Hill Inc) 288 p
[4] Kshirasagar N and Priyadarshi T 2008 *Software testing and quality assurance: theory and practice* (US: A John Wiley & Sons) 616 p
[5] Kitchenham B and Pfleeger S L 1996 *IEEE Software January* 12–21
[6] Lewis W E 2000 *Software testing and continuous quality improvement* (US: Auerbach Publishers) 620 p
[7] Galimova E Y and Kovalenko A N 2016 *Vestnik of Don State Technical University* 4 134–140
[8] Ridene Y and Barbier F 2011 *Proc. of the 5th European Conference on Software Architecture: Companion Volume* 9:1–9:7
[9] Incki K, Ari I and Sozer H 2012 *Proc. of IEEE Sixth International Conference on Software Security and Reliability Companion* 18–22
[10] Chernikov V *DevOps in the service of man: We automate the assembly, distribution and testing of new versions of the application*. Available at: https://xakep.ru/2017/02/17/devops-authomatization/
[11] Kumar V 2016 *Mobile Application Penetration Testing* (Birmingham: Packt Publishing) p 672
[12] Kumar V 2014 Smoke testing *International Journal of Scientific and Research Publications* 4 (2) 57–62
[13] Nikeshin A V, Pakulin N V and Shnitman V Z 2012 *Trudy ISP RAN* 23 387–404
[14] Grebenyuk V M 2013 *Internet-zhurnal «NAUKOVEDENIE»* 13
[15] Jiayang Liu, Lin Zhong, Jehan Wickramasuriya and Venu Vasudevan 2009 *Pervasive and Mobile Computing* 5 (6) 657–675
[16] Kazansky N, Lysuk P and Nemcov Y 2018 *Proc. of XV St. Petersburg International Conference Regional informatics “RI-2018”* 314–316
[17] Skachinsky E 2013 *Wireless Technologies* 3 32–340
[18] Jog Y, Thenuan P, Khanna D and Chavan A 2016 *Indian Journal of Scientific Research* 5 (4) 188–193
[19] Saaty T L 1985 *Analytical Planning: The Organization of Systems* (New York: Pergamon Press) 208 p
[20] Saaty T L 1987 *Decision Science* 18(2) 157–177
[21] Davnis V V 2005 *Predictive models of expert preferences: monograph* (Voronezh: Voronezh State University Publishing House) 248 p
[22] Galimova E Y 2020 *E3S Web of Conferences* 177 02008