White paper
An innovative tool to benchmark Smartphone Operating Systems

Jakajima, Nuenen, The Netherlands

This is a publication of Jakajima BV, Nuenen, The Netherlands

Table of Contents

Introduction

Jakajima, a knowledge and information provider for the telecom, ICT, and consumer-electronics trade chain, noticed that no real benchmark has been developed for Smartphone Operating Systems, while the importance of the smartphone and its' OS is growing rapidly.
Together with Eindhoven University of Technology, Jakajima has developed a benchmark for these Operating systems.

With this benchmark, Jakajima and Eindhoven University of Technology offer the market (OS developers, application developers, hardware suppliers, operators etc.) new innovative tools to compare their mobile products and/or applications in an independent way.
The benchmark is an ideal tool in any phase of a product; during concept, development and deployment.

In this white paper, based upon the research by Eindhoven University of Technology, you will find the approach, methodology and some results.

We at Jakajima, are sure that this tool will help the market to improve product/service development, so that the enduser will be able to use the products/services in the way he wants it; simple and straightforward.

We thank Vodafone and Iqua for their support.

Nuenen, the Netherlands, January 2008

1 Description of the Smartphone Operating System Benchmark

This report describes the smartphone operating system benchmark, which is conducted by Jakajima B.V. and the Eindhoven University of Technology. The experiment is carried out according to the High Contrast Consumer Test (HCCT). The HCCT is a method for testing product - user interaction in the first use phases in order to provoke product – user interaction problems in the product development process that would normally occur in the field after market introduction. This is achieved by observing extreme users when using an innovative product under realistic operating conditions. Extreme users are groups of users who use the product in an extreme way. In this context product – user interaction problems are defined as:

All problems that users experience because of the (lack of) functionality of the product

The set-up of the adapted HCCT framework consists of the following steps:

1. Identification of all innovative product features. These innovative product features need to be ranked according to feature importance by the future users of the product. The most important innovative product features need to be tested.
2. Identification of the extreme users of both the product and the features that are going to be tested. The users need to be classified into the categories, where knowledge is the contrast factor. The extreme users can be identified based on this classification.
3. Defining task scenarios. Realistic task scenarios have to be developed, which enable the observers to collect the right data. These task scenarios have to focus on the first use phases and should give the user enough freedom in the way he executes the task.
4. Defining data collection method. Specific measures for efficiency, effectiveness and satisfaction have to be defined. Per measure it needs to be decided how data on this measure is going to be collected.
5. Setting up a test session in which the extreme users can be observed while utilizing the test product under realistic operating conditions. This test should focus on the first use phases, namely the unpacking and installation process and the first use phase. To keep the test realistic no help should be provided during these tests, except when a defective product is delivered or the user wants to call the help desk / service centre.
6. Thinking aloud. To gain additional information about the user's expectations and feelings about the product they are encouraged to think aloud. This means that they say aloud what they are thinking while they use the product.
7. Data analysis and problem identification. The data analysis consists of tabulating and summarizing the quantitative data, searching for trends and surprises in the quantitative data, examining the quantitative data for problems and prioritizing all problems by severity and scope.

Each of these steps is discussed in a separate chapter in this document. First the goal of the smartphone operating system benchmark is described in chapter 2. The identification of the features for the experiment is described in chapter 3. The task scenarios are defined in chapter 4. The specific measures that are used to gather data during the experiment are identified in chapter 5. The resulting test procedure is described in chapter 6. Finally, the data analysis and resulting conclusions are given in chapter 7.

2 Goal of the Smartphone Operating System Benchmark

The goal of the smartphone operating system benchmark is to compare the performance of the major smartphone operating systems. An operating system can be defined as:

The system that controls the execution of all other programs, communication with peripheral devices and use of memory and resources (Alter, 2002).

The performance of the smartphone operating systems is determined by measuring how efficient and effective the features on each of the smartphone operating systems are and how satisfied the participants are with each of the smartphone operating systems. In this report, a feature is defined as:

A characteristic of a product, something the product has or is able to do (Gielen, 2007).

The measures from each smartphone operating system are combined, which results in an overview of the performance of the major smartphone operating systems during the first use phases based on the features that have been tested during the HCCT.

The most important smartphone operating systems are described below, as well as their market share in the third quarter of 2006 ("Symbian Fast Facts", 2007).

- Symbian (both Symbian S60 and Symbian UIQ) 72.8%
- Linux OS 16.7%
- Windows Mobile 5.6%
- JAVA 2.8%
- Palm OS 1.8%
- Others 0.3%

All of the above-mentioned smartphone operating systems, except Linux OS, have been tested according to the adapted HCCT framework. It was not possible to test Linux OS, because Linux did not participate in the experiment.
During the tests, Apple introduced the iPhone. There was not enough time to add the iPhone to this test.
These smartphone operating systems cannot be tested without testing the smartphones as well, because each smartphone operating system can only be used in combination with specific hardware, i.e. with specific smartphones. Consequently, different smartphones need to be tested to test the different smartphone operating systems. A smartphone is defined as:

A cellular handset using an open, commercial operating system that supports third – party applications (ABI Research, 2007).

To compensate for the influence of the smartphone on the performance of the smartphone operating system, the smartphone operating system companies were given the chance to select the smartphone on which they think their smartphone operating system performs best. This has resulted in the following combinations of smartphone operating systems and smartphones for this experiment:

- Symbian S60 - Nokia E61i
- Windows Mobile 6 - HTC S710
- JAVA - Blackberry 8800
- Palm OS - Palm Treo 680
- Symbian UIQ - Sony Ericsson P990i

These smartphones are shown on the cover of this report.

3 Selection of the Features for the Smartphone Operating System Benchmark

The first step of the adapted HCCT framework is to determine the features that are perceived as most important by the user in making a purchase decision for a specific smartphone operating system. To determine the most important features of the smartphone operating systems, a method called ratings of functions on a pre-established list is used in the form of a questionnaire.
The basic process for carrying out a survey research contains the following steps:

- Define the goal of the survey research;
- Identify the population and sample;
- Decide how to collect replies;
- Design the questionnaire;
- Run a pilot study;
- Conduct the survey;
- Analyze the data.

Each of these steps is described in a separate section below.

3.1 Goal of the Feature Importance Questionnaire

The first step of the adapted HCCT framework is to identify and rank all new innovative product features according to feature importance by the future users of the product. In this case, feature importance is defined as:

The relative importance of a feature, in proportion to the other features of a product

To get a clear overview of which smartphone operating system features are perceived as most important in the purchase decision, not only the new innovative features, but all features, are included in the feature importance questionnaire. Consequently, the goal of this survey is to rank all features of the smartphone operating systems based on feature importance.

3.2 Population and Sample

The population contains all members of the group that are interesting for obtaining data about the goal of the survey. In this case the population consists of all professional smartphone users in the Netherlands and Belgium. A professional smartphone user is in this report defined as:

A person who makes the decision to purchase a certain type of smartphone for his company, a person who uses a smartphone for business purposes or a person who needs to apply his knowledge of smartphones in his profession.

It has been decided to choose professional smartphone users as the population, because the majority of all smartphone users use their smartphone for business purposes.

Not all professional smartphone users can be tested, because the population is too large and it is also practically impossible to reach all professional smartphone users. Therefore a sample has been chosen, which consists of professional smartphone users.
Since there is no previous data available for the feature importance questionnaire, 200 to 300 respondents are required for this questionnaire based.

Jakajima publishes AutoConnect (www.autoconnect.nl) and Connexie Belgium (www.connexie.be), two trade magazines, which focus on resellers in telecom, ICT and the consumer electronics market. As it is likely that a reasonable amount of the subscribers to AutoConnect and Connexie are professional smartphone users, they have been chosen as sample. This means that 6000 potential professional smartphone users have been asked to participate in the survey.

3.3 Data Collection Method

The data collection method to determine the most important smartphone operating system features in the purchase decision is a technique called ratings of functions on a pre-established list in the form of a questionnaire to determine the relative importance of the features of an innovative television.

In this case an e-mail survey containing a link to a web page has been used. An e-mail is sent to all possible respondents containing the goal of the questionnaire, a brief introduction to the research and a link to the questionnaire, which is on the website of Jakajima.

3.4 Design of the Feature Importance Questionnaire

A smartphone operating system is a complex product and therefore contains more than 40 different features. If all these features are included in the questionnaire separately, this would result in a long questionnaire. This is likely to result in a high non – response rate. It has therefore been decided to group features that are used for similar activities. This grouping has resulted in a list of 23 smartphone operating system features. All smartphone operating features are deducted from the websites of the major smartphone operating system companies.
A number of general questions have been added to the questionnaire to get more information on the background of the respondent and to find out if the respondent is a professional smartphone user.

3.5 Pilot Study

A pilot test has been conducted to test if the questionnaire still contained errors, ambiguous questions or other flaws. For this pilot study one employee, one PhD student and five Master students from the Eindhoven University of Technology have been asked to fill in the questionnaire. Furthermore, the questionnaire is sent to experts of the participating smartphone operating system companies, which has resulted in one reply. The pilot study revealed some inconsistencies in question formulations and some ambiguous or unclear sentences. After adjusting the questionnaire, it has been sent to the respondents.

3.6 Survey Response Rate

The questionnaire has been distributed to approximately 6000 subscribers of the trade magazines AutoConnect and AutoConnexie in the Netherlands and Belgium. In total 494 questionnaires have been filled in and returned of which 151 questionnaires have been excluded from the analysis, because they were not filled in completely, had inconsistencies in the answers or contained scores that are not possible. This has resulted in 343 questionnaires that are appropriate for further analysis, which is enough based on the minimal required response rate determined in section 3.2. The average age of these respondents is 39 years and the majority is male (289 males and 54 females). Most of these respondents, namely 284, have a smartphone, either for business or private use. From the 284 people, 236 respondents can be classified as a professional smartphone user.

3.7 Data Analysis

The features with the highest ranks are the most important smartphone operating system features in making a purchase decision according to the professional smartphone users. The results are given in table 1 on the next page. They show that the professional smartphone users perceive certain smartphone operating system features as more important than others in making a purchase decision. The features for the smartphone operating system benchmark are selected from the Friedman test results based on the following criteria:

- The feature should have a high ranking in the feature importance measurement.
- The feature should be a typical smartphone feature.
- It should be practically feasible to test the feature in the test location.

The features network connection and personal information management (PIM) are selected, because they are the only two features that satisfy all three criteria. Furthermore, the feature making phone calls is selected although it is not an innovative feature. However, it is the most important feature according to the professional smartphone users and none of the other important and innovative smartphone features can be tested in the test facilities.

As stated in chapter 2 each smartphone operating system can only be tested in combination with specific smartphones. These smartphones do not all have the same input method (e.g. only the Palm Treo 680 and the SonyEricsson P990i have a touch screen). As these different input methods may influence the performance of specific features, it is decided to measure the influence of the input method on the performance of these features as well. However, the input method cannot be tested

separately from the other features, because it is needed to use each feature. The input method can therefore not be tested in the same way as the other features. Consequently, it has been decided to measure only the participants' level of satisfaction with the input method for the features they are testing.

Table 1: Friedman test result (features are sorted descending based on importance)

Rank	Name Feature	Mean Rank
1	Making phone calls	21,10
2	Sending messages	18,21
3	Synchronization	17,54
4	Network connection	17,33
5	Personal Information Management	17,25
6	Wireless connection	16,92
7	Security	16,89
8	Input method	16,14
9	Internet access	15,84
10	Push mail	13,97
11	User interface	12,64
12	Navigation	12,15
13	Wired connection	12,00
14	Downloading and installing applications of 3rd parties	10,69
15	VoIP	9,62
16	Media player	9,29
17	Use of applications	9,15
18	Sound options	9,06
19	Camera options	8,89
20	Voice recognition / recognition of handwriting	8,57
21	Business applications compatibility and integration	8,20
22	Recording options	6,61
23	Digital video broadcast	6,44
24	Entertainment software	5,49

4 Task Scenarios for the Smartphone operating system benchmark

The third step of the adapted HCCT framework is to define task scenarios. The features making phone call, network connection and PIM are selected for the experiment based on the feature importance measurement and the specified conditions, as described in section 3.7. A task scenario is developed for each of these features. These individual task scenarios are linked by an overall scenario. These task scenarios make the tasks more realistic and enable the observers to collect the right data about the features during the experiment. They are used in a varying sequence during the experiment to prevent ordering effects.

4.1 example Task 1 for making phonecall

You want to test the call function of the smart phone. To do this, you want to make a phone call to the telephone, which is next to the TV. The telephone number of this telephone is saved in the contact list of your smart phone as 'telefoon lab'.

For this task we ask you to call the telephone number of the lab

Fill in the corresponding questionnaire after executing this task

After this, go back to the main menu of the smart phone by pushing the red button

4.2 example Task 2 for network connection

At the moment you only have the smart phone available. For an appointment tomorrow morning at the university you want to look up the address of the TU/e on the Internet.

For this task we ask you to look up the address of the TU/e. This address can be found on the website http://www.tue.nl/nl/contact/

Fill in the corresponding questionnaire after executing this task

After this, go back to the main menu of the smart phone by pushing the red button

4.3 example Task 3 for using PIM

Today you have made an appointment with Mr. Jansen, who is your supervisor, for tomorrow from 14.00 to 15.00 o'clock to discuss the progress of your project. This appointment is not in your smart phone or on your laptop yet.

For this task we ask you to put the above-mentioned appointment in the smart phone and place a reminder for 15 minutes before the start of this appointment

Fill in the corresponding questionnaire after executing this task

After this, go back to the main menu of the smart phone by pushing the red button

Fill in the corresponding questionnaire after executing this task

Picture 1 Testroom

Picture 2 Controlroom

Picture 3 During test

5 Identification of Measures for Collecting the Quantitative Data

The adapted HCCT framework states that specific measures for measuring the effectiveness, efficiency and satisfaction have to be formulated, to enable the test team to collect the necessary quantitative data during the actual HCCT.

Efficiency
Efficiency refers to the resources expended in relation to the accuracy and completeness with which the participant achieves the specified goals (ISO, 1999). Ziefle uses three efficiency measures in her usability study about mobile phones, namely task completion time, the number of returns to a higher level in menu hierarchy and the number of detour steps. The first two measures can also be used for the experiment. The number of detour steps, which refers to the number of unnecessary key strokes, cannot be measured as this requires keystroke logging software, which is not available for this experiment. From the list of specific usability measures developed by Hornbaek (2006) it appears that the number of menus that the participant needs to complete a task provides similar data as the number of detour steps.

The three efficiency measures that are identified above are used in the smartphone operating system benchmark. They are discussed in detail below.

- Task completion time

The task completion time is defined as the time the participant needs to complete a task. The task starts when the participant starts reading the task scenario and ends if the task is successfully completed. The task completion time can be calculated by subtracting the time at the beginning of the task from the time at the end of the task.

- The number of returns to a higher level in menu hierarchy

The number of returns to a higher level in menu hierarchy is defined as the number of times a participant goes to a known position higher in the menu hierarchy. This indicates that the participant believes he has taken the wrong path and goes back to a known position in the menu in order to reorient himself. The number of returns to a higher level in menu hierarchy is calculated by counting the number of times the participant returns to a previous menu higher in the menu hierarchy.

- The number of menus used

This is defined as the number of menus the participant needs in order to successfully complete a task. A menu is in this context defined as an application on the smartphone or a hierarchical level in an application on the smartphone. This measure provides an indication of the clarity of the menu structure and is calculated by counting the number of menus a participant uses to successfully complete a task.

Effectiveness
Effectiveness refers to the accuracy and completeness with which the participant achieves the specified goals. It is therefore decided to use the number of successfully completed tasks to measure effectiveness in the smartphone operating system benchmark as well. This measure is discussed in more detail below.

- Number of completed tasks

The number of completed tasks is defined as the number of tasks that the participant completes successfully. Successfully completed does in this context mean that the tasks are executed completely and accurately. The number of completed tasks is calculated by counting the number of tasks that are successfully completed.

Satisfaction
Satisfaction refers to the freedom from discomfort and a positive attitude to the use of the product.
The interface is defined as:

Those aspects of the system that the user encounters
However, in the smartphone operating system benchmark only three features of the smartphone operating system (which is a part of the interface) are tested in a relatively short time frame. It is therefore not likely that participants can determine their level of satisfaction with the entire smartphone operating system. Consequently, it has been decided to measure the participants' level of satisfaction

with each of the features that are tested during the experiment. This measure is discussed in more detail below.

- Participant's level of satisfaction with specific features

The participant's level of satisfaction with specific features is defined as the level of satisfaction the participant experiences in using a specific feature of the product. A post - task questionnaire can be used to measure the participant's level of satisfaction with specific features. For the smartphone operating system benchmark, the After Scenario Questionnaire (ASQ) has been selected as the post - task questionnaire. The ASQ is a three-item questionnaire, which is specifically developed for scenario-based tasks. It measures the three important components of user satisfaction, which are ease of task completion, satisfaction with the time needed to complete a task and satisfaction with the support information. One question is added to the original ASQ to determine the participant's level of satisfaction with the input method of the smartphone, as discussed in section 3.7. The resulting ASQ is given in appendix 11.

Efficiency and effectiveness can be measured by using observation. Observation in this situation means that the participants are observed without intervention while they execute the predefined tasks. This means that they are free in the way they solve problems and that the observers are not allowed to steer the participants, except if they explicitly state that they cannot complete the task without help. Satisfaction measures the perceptions, opinions and judgements of participants. These measures are mainly subjective, which makes observation an inappropriate data collection method. The adapted HCCT framework therefore states that think aloud can be used to collect data on the subjective satisfaction measures. This can be achieved by collecting the participants' comments on the product.

6 Test Procedure

The test procedure briefly discusses how the experiments are executed. Before the actual experiments started, the initial test procedure has been tested in a pilot study. The goal of this pilot study was twofold. On the one hand was it an opportunity for the test team to get familiar with the test procedure. On the other hand was it a possibility to determine how the test procedure works in practice. For this pilot study, two participants with a high knowledge level and one participant with a low knowledge level have executed the experiment according to the initial test procedure. The initial test procedure has been adapted based on the outcomes of the pilot test, which has resulted in the final test procedure. This test procedure is briefly discussed below.

Each experiment started with an introduction in which the goal of the experiment was explained and the test environment was introduced to the participant. The participant also received a short written introduction about the distinguishing functionalities of a smartphone and the organizational aspects of the experiment. This introduction ensured that each participant had the same start level for the experiment. The participant had the opportunity to ask any remaining questions after reading this introduction. The actual experiment consisted of three tasks. All tasks focused on the first use phases and the participant had to install or configure the software where necessary. The participant was allowed to use all information provided in the box of the smartphone, which included the manual. Help has only been given if the participant explicitly stated that he could not complete the task on his own. This help represented the use of a helpdesk or acquaintance of the participant and was used to keep the experiment realistic and to reduce the possibility that the participant was not able to complete the task within the available time. If the participant still exceeded the time limit for a task, he was interrupted and asked to continue with the next task. After each task, the participant was asked to fill in the ASQ, which is given in appendix 11. Each participant received a token of appreciation at the end of the experiment.

Each participant has tested one smartphone operating system and the smartphone is put back into the initial settings after each experiment to ensure that every participant has the same start situation. The whole experiment took between 20 and 40 minutes, depending on the working speed of the participant and all experiments have been videotaped. During the experiment, all observations and comments from both the participant and the test team have been noted by the observers. Furthermore, data has been gathered on the measures that are defined in chapter 5. This data serves as the input for the data analysis.

7 Data Analysis and Conclusions on the Smartphone operating system benchmark

In this chapter a comparison is made for each smartphone operating system feature separately. Based on the outcomes of these analyses a general conclusion is formulated.

The results of the analyses are presented per feature in the following way:

- First a table is given containing all the quantitative data.
- Then graphs are given in which the values of all smartphone operating systems for a specific efficiency, effectiveness or satisfaction measure are shown. The significant differences for this specific measure are given below the graph.
- Finally it is concluded which smartphone operating system is most suitable for using the specific feature based on the significant differences among the smartphone operating systems for all efficiency, effectiveness and satisfaction measures.

As last part of the quantitative analysis, an overall conclusion is formulated in which the conclusions for all separate features are used to make inferences about the smartphone operating systems in general. These conclusions are only based on the quantitative data. In addition to this quantitative comparison, the problem lists are available on request per feature for all smartphone operating systems. The cause of the problems of the problem list can be in the product – user interaction, the smartphone operating system, the hardware of the smartphone or in a combination of them. No distinction is made among problems in each of these categories. The problems are grouped and sorted according to the problem groups they belong to. This makes it easier to compare the problems among the smartphone operating systems.

When interpreting the conclusions it is important to keep the following limitations of this research in mind:

- Only the first use phases are tested in the experiment. The performance of a participant over the entire lifetime of the smartphone may deviate from the performance in these first use phases.
- All the standard settings have already been placed in the smartphone before the test, which means that the participant does not have to start up the smartphone himself and therefore does not get possible welcome messages.
- Only three features of the smartphone have been tested, while a smartphone has more than 40 different features.
- The sample size was relatively small and consisted entirely of students. It is therefore not possible to generalize the results to other groups of participants.
- The students are not selected randomly. It is therefore not possible to generalize the results to all students.
- The test has been conducted in a laboratory environment. This artificial environment may have influenced the behaviour of participants.
- The participant does not get an extensive explanation about the specific smartphone he is going to test it in the experiment. He may therefore not be aware of certain functionalities of the smartphone.
- In the problem lists no distinction has been made between problems which are related to the product – user interaction, the smartphone operating system, the hardware of the smartphone or in a combination of them.

Below, the results of the analysis of all smartphone operating systems together per feature are given as well as the resulting conclusions.

7.1 Results; feature making phone call

Table 10: Comparison of the smartphone operating systems for the feature 'making phone call'

	Symbian S60	Windows Mobile 6	JAVA	Palm OS	Symbian UIQ
Task Completion Time (seconds)	65,57	61,40	87,57	105,92	58,91
Number of menus used	2,31*	3,20	4,14	10,77	3,55
Number of returns to a higher level in menu hierarchy	0,31*	0,73	1,21	1,08	0,73
Satisfaction	5,96	6,25	5,08	5,15	6,20
Percentage of participants that completed the task successfully	100	100	100	92,86	100

* This score is based on 13 participants, because for one participant it is not possible to calculate this value

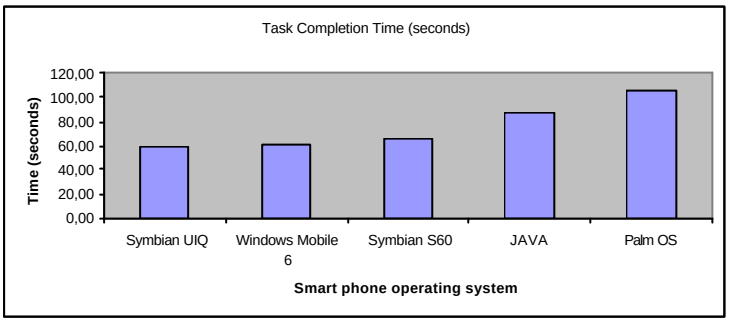

Figure 1: Comparison of the task completion times for the task 'making phone call'

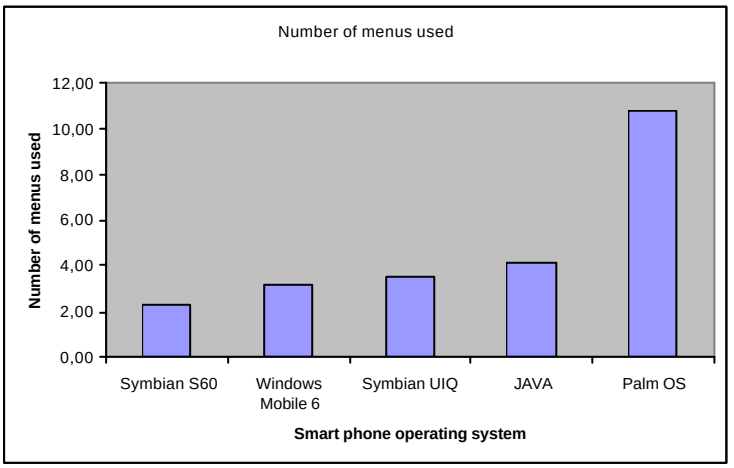

Figure 2: Comparison of the number of menus used for the task 'making phone call'

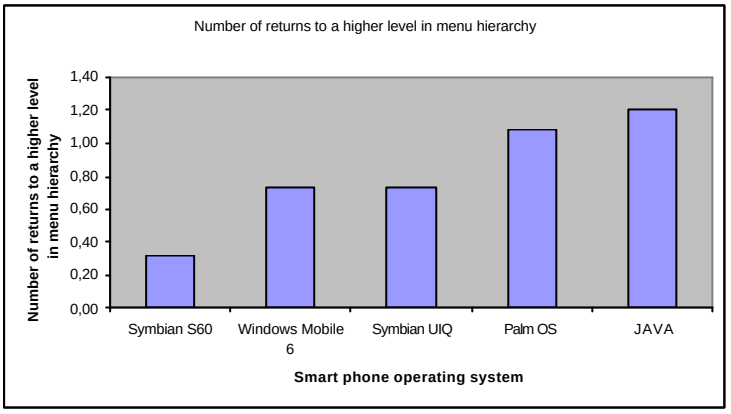

Figure 3: Comparison of the number of returns to a higher level in menu hierarchy for the task 'making phone call'

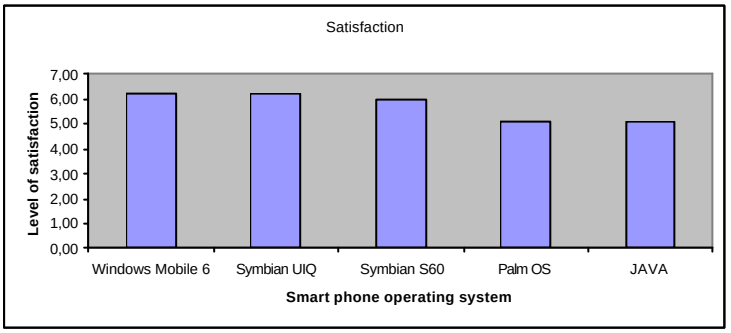

Figure 4: Comparison of the level of satisfaction for the task 'making phone call'

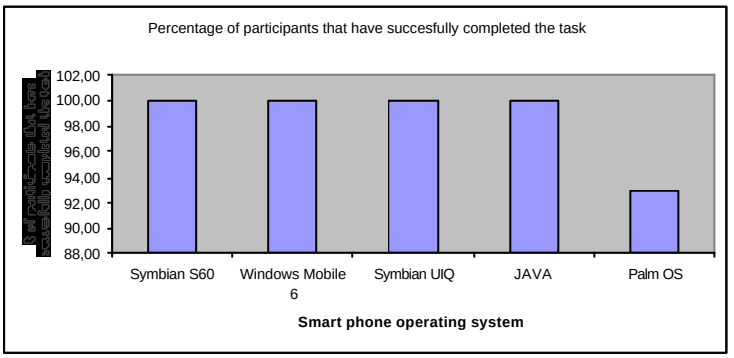

Figure 5: Comparison of percentage of participants that is able to successfully complete the task 'making phone call'

Conclusion:
It appears that none of the smartphone operating systems performs best or worst for making a phone call, because none of the smartphone operating systems performs best or worst on (almost) all measures. It does appear, however, that the Palm and JAVA operating systems perform the worst on all measures, but it needs to be noticed that the differences with the other smartphone operating systems are small.

7.2 Results; feature network connection

Table 11: Comparison of the smartphone operating systems for the feature 'network connection'

	Symbian S60	Windows Mobile 6	JAVA	Palm OS	Symbian UIQ
Task Completion Time (seconds)	195,86	400,86	552,40	239,93	427,20
Number of menus used	7,23*	7,00	23,20	1,50	24,00
Number of returns to a higher level in menu hierarchy	1,69*	1,93	8,40	0,21	9,60
Satisfaction	5,34	4,34	3,15	5,21	3,55
Percentage of participants that completed the task successfully	100	93,33	35,71	100	45,45

* This score is based on 13 participants, because for one participant it is not possible to calculate this value

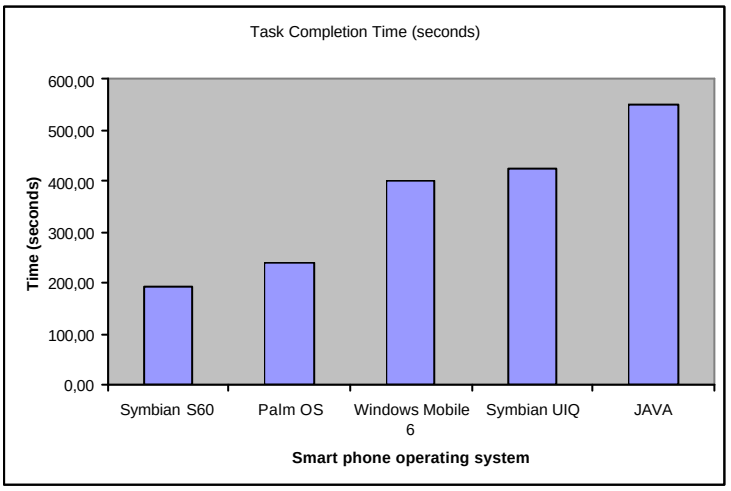

Figure 6: Comparison of the task completion times for the task 'network connection'

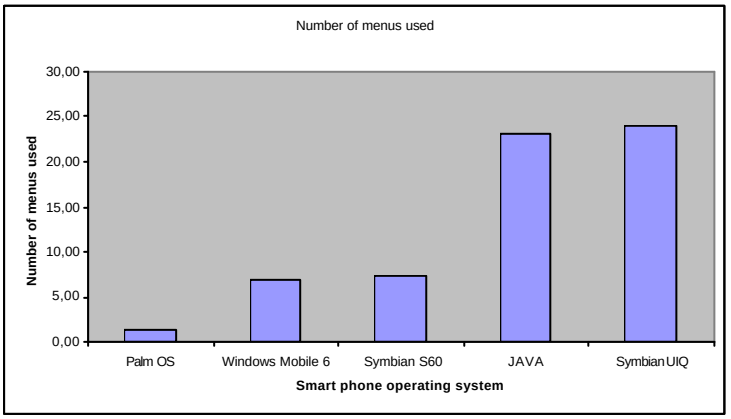

Figure 7: Comparison of the number of menus used for the task 'network connection'

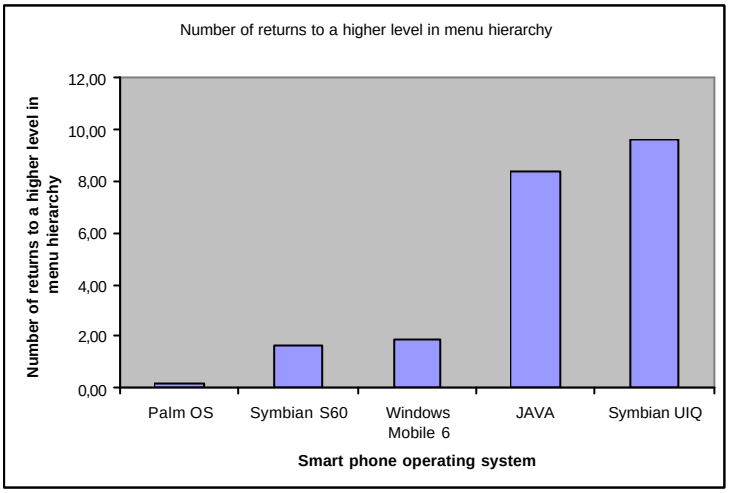

Figure 8: Comparison of the number of returns to a higher level in menu hierarchy for the task 'network connection'

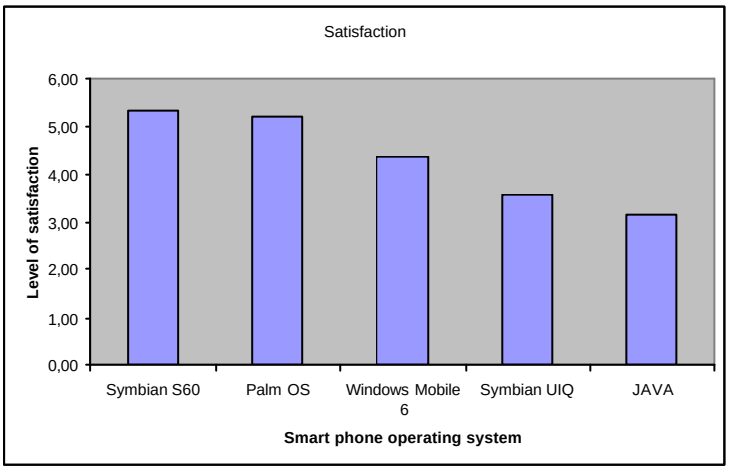

Figure 9: Comparison of the level of satisfaction for the task 'network connection'

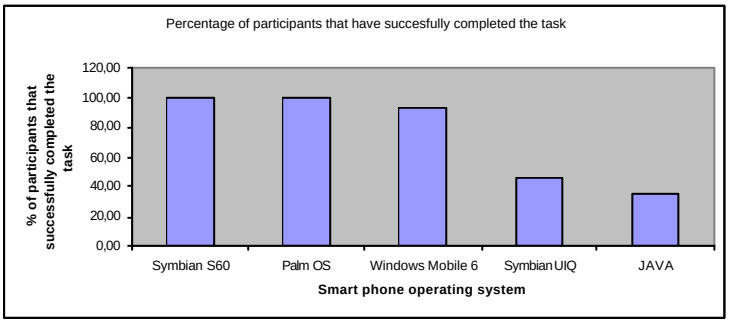

Figure 10: Comparison of percentage of participants that is able to successfully complete the task 'network connection'

Conclusion:
It can be concluded that the Symbian S60 and Palm operating system perform better than the Symbian UIQ and JAVA operating system. The Windows Mobile 6 operating system also performs better than the Symbian UIQ and JAVA operating system on all measures except the time needed to complete the task.

7.3 Results; feature PIM

Table 12: Comparison of the smartphone operating systems for the feature 'PIM'

	Symbian S60	Windows Mobile 6	JAVA	Palm OS	Symbian UIQ
Task Completion Time (seconds)	190,92	241,64	199,21	254,50	280,70
Number of menus used	9,09*	13,79	9,71	6,80	8,20
Number of returns to a higher level in menu hierarchy	1,55*	3,21	2,29	1,00	1,40
Satisfaction	5,90	5,02	5,57	4,95	5,40
Percentage of participants that completed the task successfully	92,86	93,33	100	71,43	90,90

* This score is based on 12 participants, because for one participant it is not possible to calculate this value

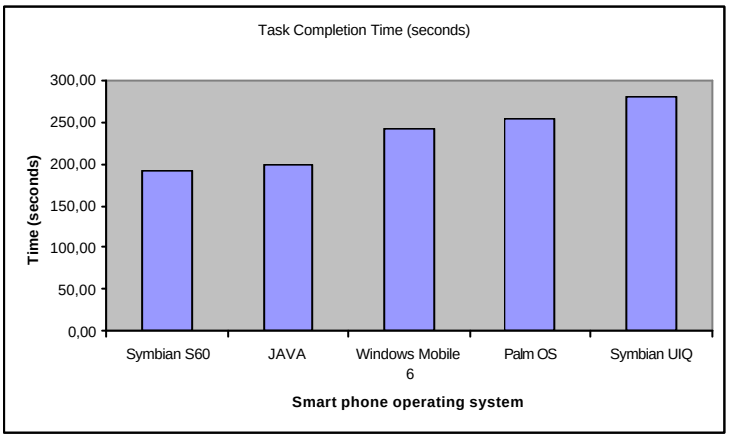

Figure 11: Comparison of the task completion times for the task 'PIM'

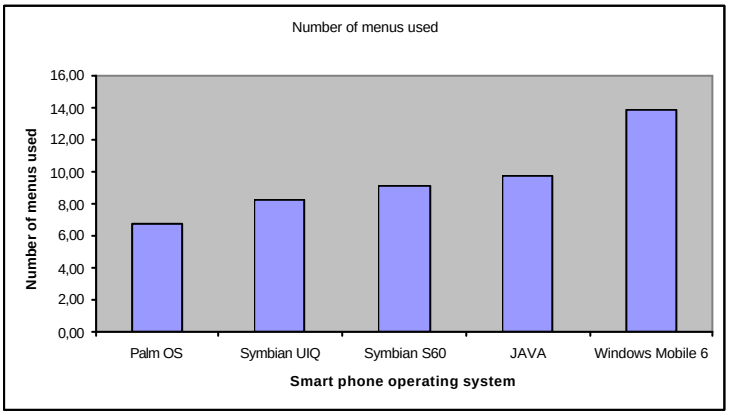

Figure 12: Comparison of the number of menus used for the task 'PIM'

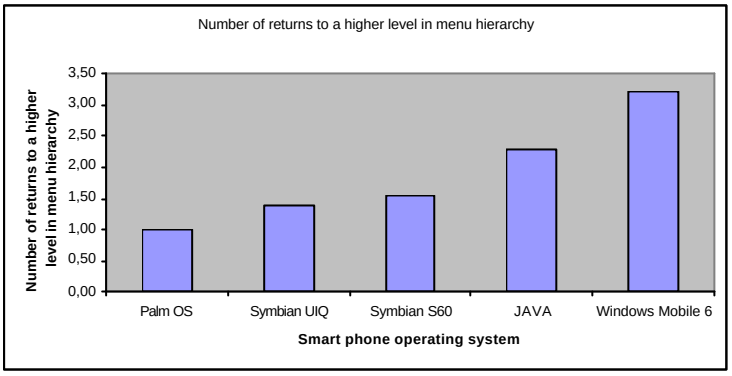

Figure 13: Comparison of number of returns to a higher level in menu hierarchy for the task 'PIM'

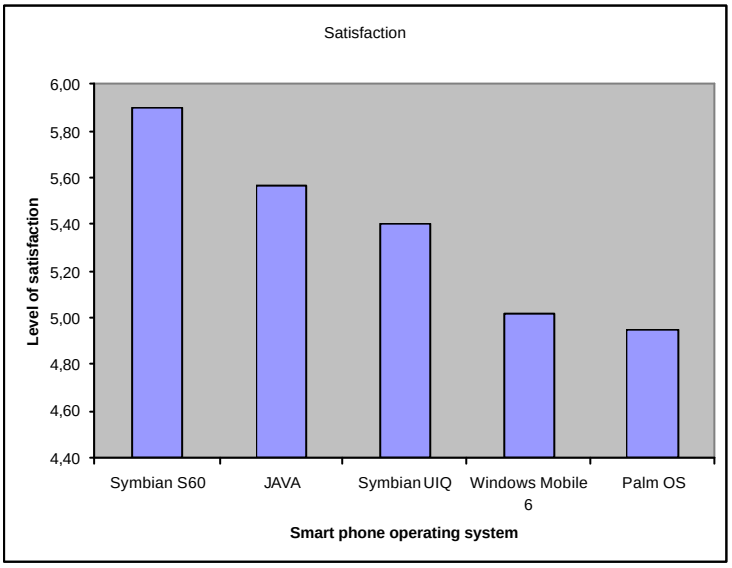

Figure 14: Comparison of the level of satisfaction for the task 'PIM'

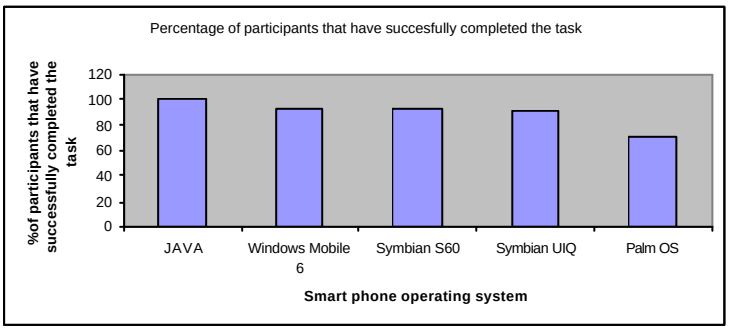

Figure 15: Comparison of percentage of participants that is able to successfully complete the task 'PIM'

Conclusion:
As can be seen in the table and the graphs, there is no consistent pattern among the different smartphone operating systems for this task. It appears that the differences are relatively small and that the smartphone operating system that performs best on one measure, performs almost worst on another measure. It can therefore be concluded that none of the smartphone operating systems performs better than the other smartphone operating systems for using the PIM functionalities, even

though the PIM functionalities are implemented in different ways on the different smartphone operating systems.

7.3 General Conclusion about the Smartphone operating system benchmark

When the conclusions for each of the tasks are combined, it appears that there is no smartphone operating system which performs best on each task. It appears that a smartphone operating system can perform best on a specific task, but perform worst on another task. It can therefore be concluded that the choice for a specific smartphone operating system should be based on the features that are perceived as most important by the user.

This benchmark shows that products / applications and services can be compared in a controllable and easy way.

8 More information

If you want more information about this benchmark, and how it could be used for other products, applications and or services, contact Jakajima.

Jakajima BV
De Pinckart 54
5674 CC Nuenen
The Netherlands

Telephone	+31 (0)40 2631131
Fax	+31 (0)84 7244694
E-mail	info@jakajima.eu
Web	www.jakajima.eu