Statistical Analysis of Mobile Processor Frequency Scaling Algorithms

Sidartha Carvalho, Pedro Fernandez and Jeronimo Junior

Abstract—Recently can be observed a growth in mobile device sales, including smartphones and tablets, with its high quantity and diversity of hardware and software models, it put on focus the fragmentation problem. Batteries with low capacity are constraints in design of embedded systems. The increase of features highlighted the need for greater capacity in mobile batteries to provide an extended time of use.

The objective of this work is statistical analysis of mobile processor frequency scaling algorithms focused on energy consumption.

We analysed the results obtained by different energy reduction techniques (the techniques employed in industry and some proposals in the literature) on the same environment. A descriptive analysis and statistical tests were done to analyze data normality. The hypothesis test about energy consumption were done with the null hypothesis that the algorithms have the same energy consumption against the alternative hypothesis that the algorithms have different energy, we used Friedman test. The Nemenyi post-test was used to define who algorithms have significant difference in energy consumption. Pearson test was realized to identify if any algorithm output have linear correlation with another.

The results reached in descriptive analysis correspond with the algorithms nature. Any significant relationship was found between the algorithms output. The 700MHz algorithm presents significant differences values compared to all other algorithms in GT-9305T case and 702MHz in GT-9505 case. The frequency near 700MHz (700MHz in GT-I9305T and 702MHz in GT-I9505) can provide a low energy consumption, but they can bring on some delay in real time applications.

1 Introduction

Devices such as smartphones and tablets that use the Android operating system grew surprisingly, increasing from 200.000 activated per day in 2010 to 1.5 million devices activated per day in 2013 and actually is the platform with more active devices [1]. This growth can be explained by the amount of features that are aggregated to smartphones and tablets each year, reducing costs of equipment and the variety of options that encompass the various user profiles, from simple and inexpensive to more complex solutions.

The need to prolong the use of devices, the strong growth of mobile devices and Android operating system, the breakthrough in the development of multicore processors, it becomes necessary to develop techniques that reduce the energy consumption of these devices to increase its time of use. This is a concern of most mobile devices design, among these are smartphones and tablets.

Devices with more features require more processing power and consequently more energy dissipation. This processing power is measured by the amount of processors and cores and the clock frequency of each core of the processor. For developing techniques that allow energy efficiency it is necessary to understand the problem addressed and the characterization of their energy spent on mobile.

The Linux operating system uses governors for operational control system. Governors are regulatory modules of the Linux kernel that communicate with the hardware. The main function is to regulate the operating frequencies of the processor cores. The Android OS is based on Linux kernel and also uses governors. There governors that focus on performance, while others, in energy savings. This choice must be made according to user usage profile [2]. Performance and energy spent minimization are opposite variables. When improve one, worse another. It is necessary to balance these variables to provide the user a better user experience. The governors more used in Linux operating system are: OnDemand, Performance and PowerSave. For Android operating system has that ones used in Linux with others adapted for mobile devices, such as Interactive, PegasusQ, UserSpace, among others.

Several techniques that reduce energy consumption in mobile devices are available in literature. You can perform energy optimizations in several areas, such as data transmission: choose intelligently and in runtime which data network is better to use, Wifi or mobile networks (2G/3G/4G) [3], decreased downtime at the end of data transmissions, this technique is called fast dormancy [4], grouping data packets and send by bursts (batching) [5] [6], decouple processing the transmission of data [4], processing off-load to the cloud, Dynamic Voltage and Frequency Scaling (DVFS) [7], among others.

The objective of this work is statistical analysis of the energy reduction techniques in mobile devices.

We describe our methodology in Section 2 and present our results in Section 3. In Section 4, we answer our three primary research questions before presenting our conclusions in Section 5.

2 Method

2.1 Questions

This study addressed the following research questions:

- RQ1 What are the characteristics of the consumption values for the selected techniques?
- RQ2 There is some sort of linear relationship between different techniques?

RQ3 Who processor frequency scaling algorithm consume less energy?

Specifically, in RQ1, we address the relationship between measures of energy consumption obtained and the particular characteristics of each algorithm. In RQ2, the hypothesis of the existence of some sort of linear relationship between the energy consumption of the different algorithms is examined. In RQ3, algorithms are compared to identify the algorithm that consume less energy consumption.

2.2 Data Acquisition

In DVFS technique is possible to reduce the operating frequency of the processor or the voltage to decrease the power expended [8]. In some cases, performance is affected by the decrease in operating frequency, while in some cases it is irrelevant. In this work the efficiency of the algorithms most used in the literature are analyzed. The energy consumption is calculated using the formula of the instantaneous power [8]:

$$p = cfv^2 + p_{static} \tag{1}$$

where the power is proportional to the frequency and voltage squared. To measure the energy consumed by the device, an energy measuring board between the smartphone and the power supply was implemented. The power supply feed the smartphone with energy while the measuring board capture and send current and voltage samples to a computer server through a mini USB connector. Scripts to automate the replication of the experiments were done and used to minimize human factor in energy consumption. In the smartphone we have a lot of dynamic variables inside the operational systems like: GPS, Data Network (WiFi or Cellular network), Calls, Applications updates, Events and the Process from the operational system itself. All the components that can be disabled will be, all the others that is impossible to disable, we will try to minimize the impact in the output variable and control that impact.

2.3 Data source

We built our database from experimentation. The experiment is composed of a set of trials and is characterized by reproduction of video stream on YouTube, playing a video during two minutes. In each trial we analyse the impacts of application of all algorithms. Each algorithm do some combination in processor frequency scaling trying to minimize energy consumption, get better performance or both. Experiments will run in smartphones: Samsung Galaxy S3 (GT-9305T) and Samsung Galaxy S4 (GT-I9505). The experiments combinations are in Table 1. The S3-700 and S4-702 represents the frequency that obtained the minimum value of energy consumption (the best case), S3-1100 and S4-1134 are the near frequency of the mean energy consumption (the mean case) and S3-1400 and S4-1890 the frequency that reached the maximum energy consumption (the worst case).

The data were stored in a .csv file with the following fields: PS, O, P, PQ, 700MHz, 1100MHz, 1400MH in case of GT-9305T device and PS, O, P, I, 702MHz, 1134MHz, 1890MHz in case of GT-I9505 device.

Table 1
Experiment Trials

	GT-9305T	GT-I9505
OnDemand	S3-OD	S4-OD
Performance	S3-PE	S4-PE
PowerSave	S3-PS	S4-PS
UserSpace	S3-US	S4-US
PegasusQ	S3-PQ	–
Interactive	–	S4-IT
700MHz Fixed	S3-700	–
702MHz Fixed	–	S4-702
1100MHz Fixed	S3-1100	–
1134MHz Fixed	–	S4-1134
1400MHz Fixed	S3-1400	–
1890MHz Fixed	–	S4-1890

Data can be found in: https://github.com/pedrodiamel/estadistica-in1119/tree/master/data. The discussion of the results is guided by the research questions. We can indicate who algorithm is the best to minimize energy consumption in the analyzed context.

3 Results

3.1 Descriptive analysis

The Fig. 1 shows an approximation of the density function for each samples studied. In a) is showing the results of the algorithms of GT-9305T and b) the result for each algorithm in GT-I9505. It can be seen that some of the samples have an apparently normal behaviour. For GT-9305T apparently the algorithms: PE, PQ, 700MHz and 1100MHz have lower consumption than algorithms: PS, OD, 1400MHz, with values below 220 Joules. For GT-I9505 the algorithms: PQ, 702MHz and OD are apparently of less consumption, with values below 250 Joules.

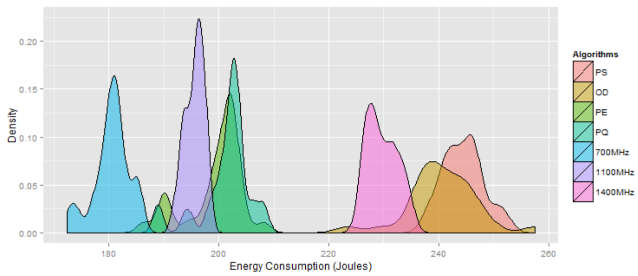
The Table 2 and 3 show a summary of the descriptive variables calculated for GT-9305T and GT-I9505 respectively. Values suggest that the algorithms 700MHz and 702MHz may be the lower consume in each case with average equal to 180.445 Joules and 230.5 Joules respectively.

Table 2
Descriptive Statistics for Samsung GT-9305T

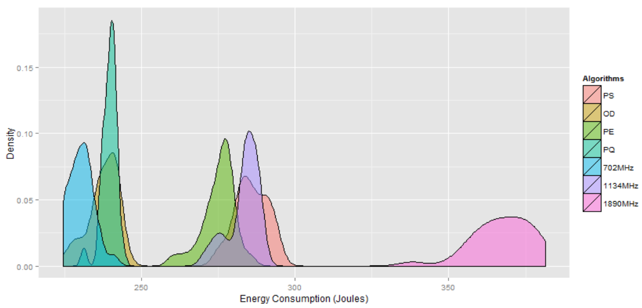
	Min	1st Qu.	Median	$Mean \pm Sd$	3st Qu.	Max
OnDemand	223.33	237.7	240.6	240.7 ± 6.22	244.1	257.5
Performance	186.4	198.7	201.2	199.3 ± 5.18	202.4	208.2
PowerSave	238.0	242.1	244.7	244.5 ± 3.64	246.5	252.2
PegasusQ	188.9	200.7	202.5	201.3 ± 4.60	203.5	208.2
700MHz	172.4	179.3	181.0	180.5 ± 3.26	182.3	185.6
1100MHz	191.8	194.1	195.9	195.6 ± 1.68	197.0	198.2
1400MHz	226.2	227.5	229.0	229.7 ± 2.67	231.8	234.5

The Fig. 2 shows a box plot graph for GT-9305T (a) and GT-I9505 (b). For GT-9305T, algorithms as OD, PE, PQ and 700MHz present several outliers and the algorithms OD and 1400MHz show some asymmetry. For GT-I9505, except the PS algorithm, all algorithms have outliers. The PE algorithm shows some asymmetry.

To confirm the hypotheses about normality from samples of each algorithm, the adherence tests for normality were

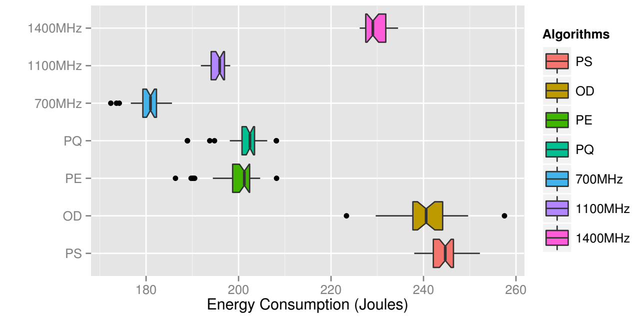


a)

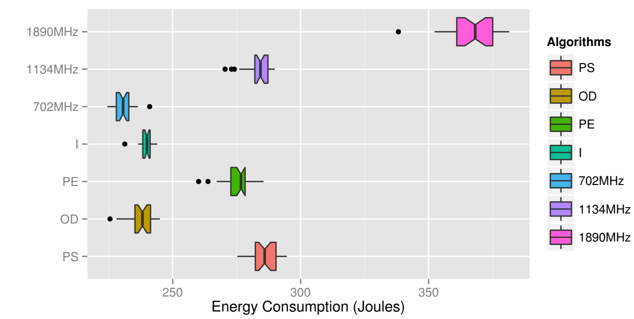


b)

Figure 1. Density approximation for each samples in GT-9305T a) and GT-I9505 b) devices.



a)



b)

Figure 2. Energy consumption Boxplots for Samsung GT-9305T and GT-I9505.

Table 3
Descriptive Statistics for Samsung GT-I9505

	Min	1st Qu.	Median	$Mean \pm Sd$	3st Qu.	Max
OnDemand	225.51	235.25	238.19	237.65 ± 4.82	241.33	245.03
Performance	260.13	272.72	276.67	275.14 ± 5.29	278.29	285.50
PowerSave	275.26	282.3	285.95	285.83 ± 5.18	290.37	294.6
Interactive	231.29	238.31	239.96	239.57 ± 2.38	241.17	244
702MHz	224.49	227.99	230.63	230.48 ± 3.97	232.85	241
1134MHz	270.48	282.09	284.3	283.30 ± 4.99	287.23	289.9
1890MHz	338.2	361	368.20	367.37 ± 9.67	375.05	381.5

done. The non-parametric Shapiro-Wilk test(SW)[9] and Lilliefors (Kolmogorov-Smirnov) test (LKS) [10] were applied. These tests are widely used in the literature for normality tests. We defined the null hypothesis H_0 that data follow a normal distribution, against the alternative hypotheses H_1 that samples do not follow a normal distribution. The p-values obtained for each of these samples are shown in Table 4 for GT-9305T and Table 5 for GT-I9505.

Table 4
Normality Tests for Samsung GT-9305T

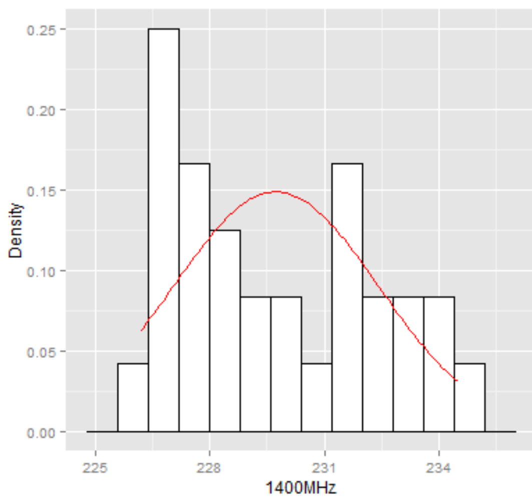
	LKS	SW
OnDemand	0.21	0.2109
Performance	0.000768	0.0009073
PowerSave	0.6164	0.6321
PegasusQ	0.01786	0.0008653
700MHz	0.195	0.09014
1100MHz	0.2782	0.1943
1400MHz	0.2689	0.02742

In GT-9305T, PE and PQ algorithms rejected the null hypothesis with significant level of 0.000768 (0.0009073) and 0.01786 (0.0008653) respectively, so there is not enough evidence to say that these samples follow a normal distribution. In the case of 1400MHz algorithm, the two test show different criteria about normality with significance of 0.05. In this case, we applied Anderson-Darling test for normality reaching a p-value of 0.05013, then the null hypothesis is rejected for a Level of significance of 0.05. This distribution also presents a kurtosis of 1.82 and Skewness of 0.34, these factors indicate a flattening and a level of asymmetry to the left of the sample distribution. Fig. 3 a) and b) present Histograms and QQ-Plots graphs from these samples. As can be seen, it reinforce the results.

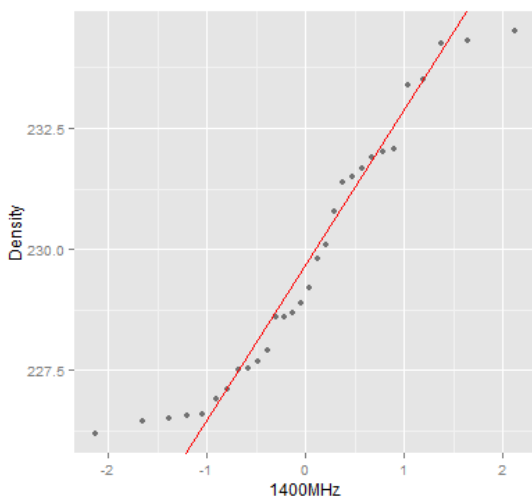
Table 5
Normality Tests for Samsung GT-I9505

	LKS	SW
OnDemand	0.08668	0.0426
Performance	0.2072	0.08294
PowerSave	0.3423	0.4645
Interactive	0.5022	0.01585
702MHz	0.9548	0.3691
1134MHz	0.01896	0.003784
1890MHz	0.652	0.1346

In GT-I9505, the 1134MHz algorithm rejects the null hypothesis with a significance level of 0.000768 (0.0009073) and 0.01786 (0.0008653) respectively, so there is not enough evidence to say that this sample follow a normal distribution. For the algorithm OD the two tests show different criteria about normality with significance of 0.05. In this case, Anderson-Darling test for normality was used again, getting a p-value of 0.44, thus we can not reject the null

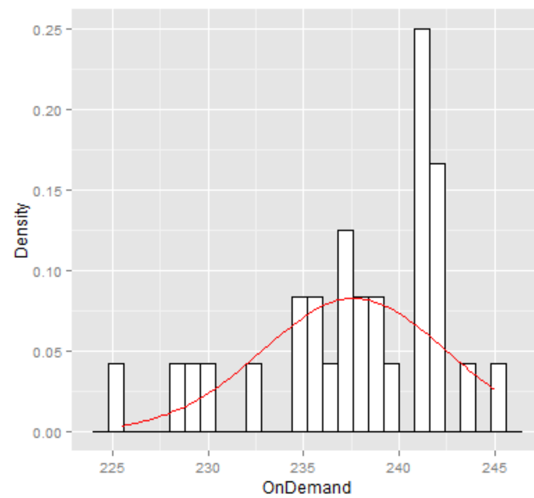


a)Histogram

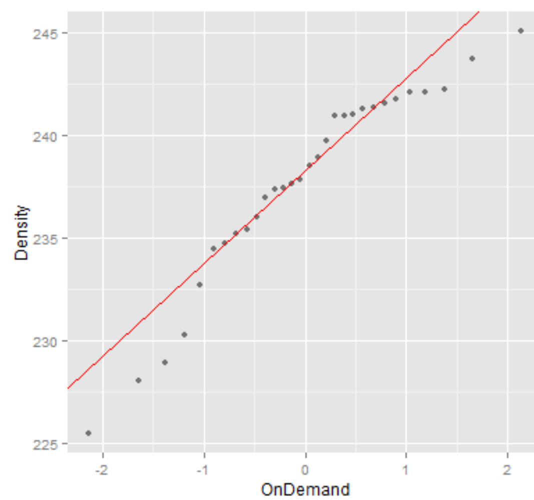


b)QQ-Plot

Figure 3. Histograms and QQ-Plots for 1400MHz algorithm in GT-9305T



a)Histogram



b)QQ-Plot

Figure 4. Histograms and QQ-Plots for OD algorithm in GT-9505

hypothesis that the outputs of the algorithm follow a normal distribution. Fig. 4 a) and b) present Histograms and QQ-Plots graphs from these samples. As can be seen, it reinforce the results.

Once the samples were characterized in the study, statistical analysis tools that will be used to answer the research questions were determined. For RQ1, a summary is made from the measured data and descriptive characteristics of normally were found in the data. To RQ2, Pearson coefficient of correlation is analyzed between the outputs of each algorithm to determine whether there is any linear relationship between the algorithms outputs, if confirmed relationship, regression tools could be used. To RQ3, the comparison between different obtained samples is performed. Because there are samples: PE, PQ and 1400MHz in the case of GT-9305T. In the case of GT-I9505, only 1134MHz do not exhibit normal distribution, nonparametric test under the following hypotheses was applied:

$$H_0 : M_1 = M_2 = ... = M_n \tag{2}$$

$$H_1 : \exists M_i, M_j \mid M_i \neq M_j \tag{3}$$

In this paper, for the hypothesis testing was selected Friedman test, the non-parametric version of Analysis of Variance (ANOVA). To determine which algorithms have significant differences, Nemenyi post-test was applied. For this analysis, only PE, PQ, 700MHz and 1100MHz in the case of GT-9305T and PE, PS, OD, IT, 702MHz and 1134MHz of GT-I9505 algorithms were used, since there are algorithms which have a much higher consumption, obviously have significant differences.

4 Discussion

In this section, we discuss the answers to our research questions.

4.1 What are the characteristics of the consumption values for the selected techniques?

The Table 6 shows that the algorithms that which were not designed to mobile devices (adapted from Linux), they do not have mobile specific characteristics and only focus on processor, for that, they have higher energy consumption. The OD algorithm has been a scaffold for PQ and IT. Can be observed that for GT-I9505 it have a lower energy consumption.

Table 6
Summary of the algorithms proposed in the literature.

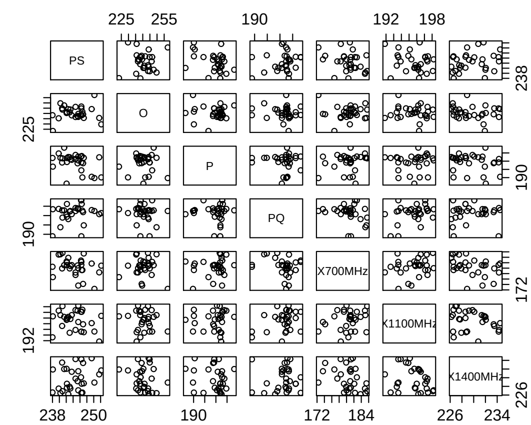
Algorithm	Description
PS	It was formulated to provide lower energy consumption, but it was not adapted to mobile devices. Mobile devices have a lot of aggregated features that consume more energy, not only the processor. It was focused in the relation between processor frequency and energy consumption, regardless the other things that consume energy. It has a high energy consumption in both scenarios (over 230 Jules) and maximum values of 252.2 and 294.6 Joules for GT-9305T and GT-9505 respectively. The average consumption is 244.5 ± 3.64 for GT-9305T and 285.83 ± 5.18 for GT-9505. No outliers.
PE	The PE algorithm was not built for mobile devices and is focused only to provide a better performance for the user. Keeping the processor frequency always in the maximum, reducing processing time and doing faster processor dependent tasks. It presents a similar behavior in both scenarios. Has an average consumption compared to other algorithms (199.3 ± 5.18 for GT-9305T and 275.14 ± 5.29 for GT-9505). Present outliers in both case.
OD	The OD algorithm was not developed for mobile devices too. It try to do frequency scaling based in processor load, keeping lower frequency when the mobile is not doing hard tasks and changing to a higher frequency when have work to do. It has an average consumption of 240.7 ± 6.22 for GT-9305T and 237.65 ± 4.82 for GT-9505.
PQ	The PQ algorithm was used for some time in Samsung smartphones, but was replaced by IT, it have a good trade off between performance and energy consumption, but have the same problem than OD algorithm. It has an average consumption of 201.3 ± 4.60 (less than OD). The minimum consumption achieved by this algorithm is 188.9 Joules.
IT	The IT algorithm was developed specifically for mobile devices and it is an evolution of OD algorithm. They differ basically in the time required to reach the optimal processor frequency, when OD change slowly, the IT change from one frequency to the best immediately. This algorithm is used in a big quantity of smartphones today, Samsung and Motorola devices. The average consumption of IT is 239.57 ± 2.38 (less than OD). This is one of the algorithms that consume less energy in this scenario.

The other frequencies, 700MHz, 702MHz, 1400MHz and so on keep the frequency without frequency scaling. Each scenario can have a optimal frequency (the best trade off between energy and performance). Algorithms PS, OD and 1400MHz in GT-9305T device, and 1890MHZ, 1134MHz, PE and PS in GT-I9505 have much higher values of energy consumption than others. A statistical analysis is not recommended when you have a large gap between the samples data.

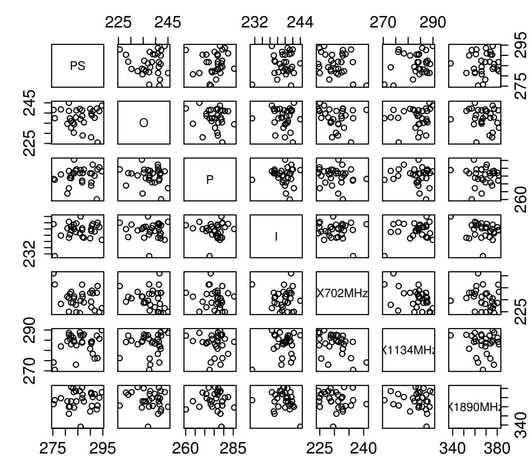
The frequency of 700MHz in GT-9305T and one of 702MHz, PQ or OD in GT-9505 probably have the lowest energy consumption.

4.2 There will be some sort of linear relationship between different techniques?

To answer this questions, the Person correlation coefficient between the output of each algorithm is analyzed to determine whether there is any linear relationship between the algorithms outputs. If this relationship is confirmed, it is possible to use regression tools. As shown in Fig. 5, apparently not significantly correlated measures are appreciated. The calculation of the correlation coefficient of Person does not show a significant relationship between the different algorithms.



a)



b)

Figure 5. Pearson Correlation Plot for GT-9305T and GT-9505

In the case of GT-9305T, the strongest relationship was found between algorithms PS and PE with reason of -0.390. In GT-I9505, the strongest relationship was found between 702MHz and 1134MHz with reason of 0.440. In general, there are no significant relationships between different algorithms.

4.3 Who processor frequency scaling algorithm consume less energy?

To detect which algorithms have the lowest energy consumption for each device, firstly, Friedman test was performed to verify if there are differences in at least two of the algorithms. This analysis discarded PS, OD and 1400MHz algorithms for the case of GT-9305T and 1890MHz for GT-I9505 because they have, of course, bigger values than other algorithms.

In the case of GT-9305T, the null hypothesis H_0 that the medians are equal is rejected with a significance level of 7.414e-15. Therefore, we can say that there are significant differences between the medians of the algorithms. Table 7 show the p-values for the Nemenyi post-test. As can be seen, the algorithm 700MHz presents significant differences values compared to all other algorithms and also has the lowest value of the average, which is demonstrated that the algorithm have strong evidences to be who have the lowest consumption.

Table 7
Nemenyi Multiple Comparison Test for Samsung Galaxy GT-9305T

	PE	PQ	700MHz
PQ	0.93208	-	-
700MHz	6.3e-11	8.4e-13	-
1100MHz	0.01436	0.00181	0.00083

In the case of GT-I9505, the null hypothesis H_0 that the means are equal, it is rejected with a significance level of 2.2e-16. Therefore, we can say that there are differences between the means of the algorithms. Table 8 shows the p-values for the Nemenyi post-test.

Table 8
Nemenyi Multiple Comparison Test for Samsung Galaxy GT-I9505

	PS	OD	PE	IT	702MHz
OD	1.3e-10	-	-	-	-
PE	0.0769	0.0005	-	-	-
IT	2.9e-08	0.9624	0.0119	-	-
702MHz	4.3e-14	0.3030	7.9e-09	0.0435	-
1134MHz	0.9985	2.1e-09	0.2033	3.4e-07	5.7e-14

As can be seen, 702MHz algorithm presents significant differences values compared to all other algorithms and also has the lowest value in average, which is demonstrated that the algorithm have strong evidences to be who have the lowest consumption. The frequency near 700MHz (700MHz in GT-I9305T and 702MHz in GT-I9505) can provide a low energy consumption, but they can bring on some delay in real time applications, it is important to have another study that analyze energy consumption and execution time to have a greater vision of this problem that affect billions of devices in the world.

The algorithms like IT, PQ and OD try to have balance in this two constraints, time and energy consumption. While PS focus only in energy reduction (which often does not work) and PE focus only in performance, without any concern about have low energy consumption.

5 Conclusion

We analyzed some processor frequency scaling algorithms for mobile devices, we used smartphones Samsung Galaxy GT-I9305T and Samsung Galaxy GT-I9505 to do the experiments.

The results reached with descriptive analyses show a correspondence with the algorithms nature, showing a higher energy consumption when using algorithms that were not built for mobile devices.

Any significance in relationship between the algorithms studied was found. In the case of GT-9305T, the strongest relationship was found between algorithms PS and PE with reason of -0.390. In GT-I9505, the strongest relationship was found between 702MHz and 1134MHz with reason of 0.440.

The 702MHz algorithm presents significant differences values compared to all other algorithms in GT-9305T case and 702MHz in GT-9505 case. The frequency near 700MHz (700MHz in GT-I9305T and 702MHz in GT-I9505) can provide a low energy consumption, but they can bring on some delay in real time applications, it is important to have another study that analyze energy consumption and execution time to have a greater vision of this problem.

References

[1] Statista, "The statistics portal," *Available in: http://www.statista.com/statistics/278305/daily-activations-of-Android-devices. Access in November - 24 - 2014*, 2014.

[2] A. S. V. Pallipadi, "The ondemand governor: past, present and future." *In Proceedings of Linux Symposium, vol. 2, pp. 223-238*, 2006.

[3] G. S. W. K. J. W. G. P. Perrucci, F. H. P. Fitzek, "On the impact of 2g and 3g networks usage for mobile phones' battery life." *in European Wireless Conference*, 2009.

[4] T. M. R. C. L. J. Y. Z. Q. L. F. Xu, Y. Liu, "Optimizing background email sync on smartphones." *In Proceeding of the 11th International Conference on Mobile Systems, Applications, and Services, MobiSys 13, pages 5568, New York, NY, USA, ACM*, 2013.

[5] A. S. R. Palit, K. Naik, "Impact of packet aggregation on energy consumption in smartphones," *In Proceedings of the 7th International Wireless Communications and Mobile Computing Conference pp. 589-594, Istanbul, Turkey*, 2012.

[6] Z. M. S. S. O. S. J. Huang, F. Qian, "Screen-off traffic characterization and optimization in 3g/4g networks," *In Internet Measurement Conference (IMC'12)*, 2012.

[7] F. Q. B. S. A. L. O. C. J. A. L. M. S. P. H. R. C. R. C. G. M. A. G. Silva-Filho, P. T. L. Bezerra, "Energy-aware technology-based dvfs mechanism for the android operating system." *In Proceedings of Brazilian Symposium on Computing System Engineering (SBESC)*, 2012.

[8] M. Domeika, "Software development for embedded multi-core systems: A practical guide using embedded intel architecture," *Newnes, Newton, MA,*, 2008.

[9] E. A. Oztuna D and T. E., "Investigation of four different normality tests in terms of type 1 error rate and power under different distributions," *Turkish Journal of Medical Sciences*, pp. 171–176, 2006.

[10] H. Lilliefors, "On the kolmogorov-smirnov test for normality with mean and variance unknown," *Journal of American Statistical Association, vol. 62, no.318, pp.*, pp. 399–402, 1967.