Exploring research hypotheses in green computing

Ikram Hamizi, Zamira Kholmatova, Giancarlo Succi

1Institute of Software Development and Engineering, Innopolis University, Innopolis 420500, Russia.

Correspondence to: Prof. Giancarlo Succi, Institute of Software Development and Engineering, Innopolis University, Innopolis 420500, Russia. E-mail: g.succi@innopolis.ru


Received: 9 Mar 2021 First Decision: 26 May 2021 Revised: 13 Jun 2021 Accepted: 13 Jul 2021 Published: 29 Jul 2021

Academic Editor: Patricia Melin Copy Editor: Xin-Jun Chen Production Editor: Xin-Jun Chen

Abstract

This paper reviews the research work done in the last 11 years in the area of green computing and analyzes the associated hypotheses, which are then structured in taxonomy and explored to pave the way for a comprehensive view of future research in the area. With the help of the taxonomy, we can understand which of the problems needs more attention. For example, when there is a small number of studies related to a problem, research needs to be conducted on this topic, while a huge number of studies could raise contradictory results that can be aggregated to a unified answer with a meta-analysis. Among the hypotheses generated, one can choose to investigate the hypotheses with a sufficient number of papers.

Keywords: Research synthesis in software engineering, meta-analysis in software engineering, evidence-based software engineering

1. INTRODUCTION

The growth of energy consumption has become a vital problem affecting everyday life. The number of smartphone users is approximately 6378 million and the number of sold laptops exceeds 220 million. This technology impacts the environment through the production, use, and disposal of digital devices. In addition,
sophisticated users also want their smartphones and laptops to work longer between charging. Thus, it is important to find a way to reduce the energy consumed by such devices. One of the existing solutions is to reduce the energy consumed by the software running on them.

The problem of energy spent by different devices has turned into the problem of understanding how to develop applications that are “energy friendly”, and this problem has already been partially addressed in the literature on the impact of development approaches \cite{1,2} and design choice \cite{3} on energy consumption. Many researchers have proposed tools and methods for estimating the energy spent by different devices \cite{4–6}. Thanks to the investigation done in this area, developers can track the energy consumed by their software, find “code smells”, and improve the efficiency of their products.

Due to the great variety of existing approaches, frameworks, and tools for developing applications, the problem of energy-efficient code can be explored from numerous perspectives. Now, we need to understand the current open questions in this area. To provide a novel approach, we decided to investigate the possible hypotheses that can be derived from the studies existing in the considered field. To do this, we firstly conducted a literature review in the area of energy consumption in software engineering. The literature review gave us a vision of the problem at the current moment. The awareness about the energy problem from the software side has already awakened, and the work on optimizing energy consumption is in progress \cite{7,8}. The second step was to select studies with empirical results to define possible hypotheses. We found 13 papers about tools for estimation of energy consumption and 39 and 35 papers related to the impact of design choice and platform issues on energy consumption, respectively.

The basic dimensions of a problem led us to the formulation of the important questions \cite{9}. In our case, these dimensions can be different devices, how energy consumption was measured, and the research topics addressed by the papers. Based on this information, we defined the following research questions:

RQ\textsubscript{1}: What are the papers that addressed energy consumption in software engineering from 2010 to 2020?
RQ\textsubscript{2}: For which devices was the energy consumption measured in the existing studies?
RQ\textsubscript{3}: What hypotheses can we derive from the studies found?

With the help of such investigation, we can understand the current trends in “green” computing, namely which algorithms, methods, and techniques are the more efficient.

This work is organized as follows. The Review Protocol section presents the protocol followed to review the existing work. The Results section presents the results of the review. The Discussion section discusses the results found. The Conclusion section draws some conclusions.

2. REVIEW PROTOCOL

Our research was conducted based on a guideline for systematic reviews appropriate for software engineering researchers proposed by Kitchenham \cite{10}. The search strategy for this investigation was defined by the research questions stated above.

To find relevant studies, we used a manual search with the Google Scholar engine since it includes most peer-reviewed scientific journals and conference proceedings. We searched for all papers related to energy consumption in software development from 2010 to 2021. The search strings were generated using the PICO approach \cite{11}. The number of papers found for each search query is presented in Table \ref{tab:1}.

To be included for the further consideration, the paper should:
Table 1. Search queries

<table>
<thead>
<tr>
<th>Search query</th>
<th>Number of papers</th>
</tr>
</thead>
<tbody>
<tr>
<td>(energy OR power) (consumption OR efficiency) in software development</td>
<td>74</td>
</tr>
<tr>
<td>(energy OR power) (consumption OR efficiency) in (mobile devices OR smartphones)</td>
<td>129</td>
</tr>
<tr>
<td>(energy OR power) (consumption OR efficiency) in (cloud systems)</td>
<td>70</td>
</tr>
<tr>
<td>(energy OR power) (consumption OR efficiency) in (embedded systems)</td>
<td>67</td>
</tr>
</tbody>
</table>

Figure 1. Comparison of the number of papers before and after applying the snowballing method.

- be peer-reviewed;
- be written in English; and
- contain information about energy consumption from the software perspective.

After quality assessment of studies based on the title, abstract, and keywords, we rejected 43.4% of papers. During the search process, we identified three big categories based on the devices for which the energy consumption was measured. These categories are mobile devices, embedded, and cloud systems.

However, after these steps, one should check the references in the already selected studies (backward snowballing) or check the studies in which the reference part contains the selected one (forward snowballing). We increased the number of papers found by using forward\cite{12} and backward\cite{13} snowballing methods. Both of them include manual and automatic search and selection of suitable studies. The comparison of the number of papers found before and after applying snowballing is presented in Figure 1.

To generate hypotheses, we need to select papers with empirical results. Based on quantitative values provided in each paper, techniques such as meta-analysis\cite{14} can help further to test hypotheses and find confounding factors.

To form new hypotheses, we defined the following criteria for paper selection:

- containing empirical results from conducted experiments;
- comparison of two or more different techniques, methods, of languages; and
- containing all presented values in joules or watts.

During the search process, we excluded duplicated and irrelevant studies. If a paper does not provide the full information about experiments done or misses the exact values, it was also removed from further consideration. Overall, after the second quality assessment, we rejected 44% of the found papers and left only 189 papers containing empirical results.
3. RESULTS

In the previous section, we describe the process of obtaining the results of data extraction from primary studies. This section reports and analyzes the findings to answer the stated research questions.

What are the papers that addressed energy consumption in software engineering from 2010 to 2020?

Figure 1 shows the 340 papers found that are related to energy consumption in software development.

For which devices was energy consumption measured in existing studies?

The available studies were divided into three groups considering the devices for which energy consumption was measured:

- mobile devices;
- cloud systems; and
- embedded systems.

What hypotheses we can derive from the studies found?

Out of the studies we found, only 189 papers contain empirical results. We analyzed them and highlighted 78 studies that can form plausible hypotheses. The rest of the papers contain results that can be explained by a direct relationship, for example, the increase of power or energy consumption with higher CPU usage.

Since we are interested in exploring research hypotheses, we further proceeded with our work using the 78 highlighted studies. During the analysis of these studies, we found that the generated hypotheses could be divided into the following groups:

- hypotheses related to the tools for estimating energy consumption (Table 2);
- hypotheses related to the impact of a design choice on the energy consumption (Table 3); and
- hypotheses related to the impact of platform-specific issues on the energy consumption (Table 4).

The first group of hypotheses compares the energy consumption values obtained from an estimation tool with the real ones. The second group describes how the design choice while developing an application can affect the energy spent by different devices. The third group addresses the relationship between platform-specific...
To better understand the structure of the generated hypotheses, we decided to represent them in a mind map (Figure 2). As described above, we have three big categories: design choice, estimation tools, and platform issues. Opposite to each category, we present the general context of available hypotheses.

### 4. DISCUSSION

The presented review aimed to probe the current possible hypotheses that can be further investigated. In the areas of energy consumption of mobile devices, embedded, and cloud systems, 340 papers were considered. Of these papers, 189 studies contained empirical results, and we were able to generate plausible hypotheses only from 78 studies. These studies were divided into three groups named: tools for estimation, design choice, and platform issues.
With the first group of hypotheses, the developers can choose the appropriate tools to estimate the energy spent by their application. The second group can help them in making informed decisions about the design choices that would lead to the production of energy-friendly software. The third group helps in detecting the platform characteristics that consume less energy while using them. The second and third groups of hypotheses are mostly under consideration while developing the software.

As shown in Figure 1, the majority of studies found during the search revolved around the energy consumption of embedded systems. Nevertheless, the resulting list of hypotheses revolves more around cloud systems and mobile devices, corresponding to the relevant papers that passed the quality assessment.

In mobile devices, the trend in the last 10 years seems to focus on proposing different offloading schemes (13 papers), a technique that migrates mobile tasks to a cloud infrastructure to reduce the local energy consumption. In a similar vein, the studies found on cloud systems focused on implementing different resource allocation techniques and studying their effect (nine papers). These techniques include VM placement techniques (VM allocation and VM migration), task and resource consolidation and allocation, and many papers focused on the dynamic voltage and frequency scaling (DVFS) task-allocation technique. As for embedded systems, papers seem to mostly address hypotheses from the first group (tools for estimation).

From all of these groups, the hypothesis considering different offloading schemes was the most studied one. The hypotheses related to the different scheduling techniques, network technologies, sorting algorithms, and programming languages also include more studies than the rest. Thus, these topics could be further investigated using different aggregation techniques to derive a general conclusion.
5. CONCLUSION

The review conducted within this project shows the not yet investigated hypotheses that can play an important role in the development of green software. With the help of the presented hypotheses tables, we can see which hypotheses need more attention. In Table 2, we can see that, for most of the hypotheses, we have only one paper. Thus, all groups related to the tools of estimation are lacking papers. This situation makes the choice of suitable tools difficult. On the one hand, it is understandable that there are many solutions for estimation and they vary from one tool or method to another. On the other hand, there is a need for a standard solution for the developers to understand how much energy is spent by their applications.[7] Hypotheses with a sufficient number of papers can be further investigated using meta-analytical techniques. Such investigation will statistically prove if one method or tool is more energy-efficient than another one.

 Needless to say, our research has some limitations. It is limited to the articles we found using the Google Scholar engine and to the scope of the search set up for the period from 2010 to 2021. It is not guaranteed that we found all relevant papers. Nonetheless, we believe that our findings are important. They define the direction for future research and show questions that are still open.

This research became a step towards understanding how to make software “green”. Further, we will consider different meta-analytical techniques that will allow us to consider how the presented hypotheses make a choice in favor of more efficient solutions. The understanding of more efficient solutions can help us to integrate them into industry and reduce the overall consumption of energy.

6. FUTURE WORK

As mentioned above, this set of hypotheses can help us to understand the direction for future investigation. From the results shown in the tables, we can understand for which hypotheses we have a sufficient number of papers and which ones can be tested.

As future work, we choose the hypotheses related to the design choices. The understanding of energy efficiency of different algorithms and methods will help us to make software “green”. In our next step, we will focus on the energy consumption of different sorting algorithms.

DECLARATIONS

Acknowledgments
This research project is carried out under the support of the Russian Science Foundation Grant № 19-19-00623.

Authors’ contributions
Made substantial contributions to conception and design of the study and performed data analysis and interpretation: Hamizi I, Kholmatova Z, Succi G
Performed data acquisition, as well as provided administrative, technical, and material support: Hamizi I, Kholmatova Z, Succi G

Availability of data and materials
Not applicable.

Financial support and sponsorship
This research project is carried out under the support of the Russian Science Foundation Grant № 19-19-00623.
Conflicts of interest
All authors declared that there are no conflicts of interest.

Ethical approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Copyright
© The Author(s) 2021.

REFERENCES


42. Park JJ, Hong JE, Lee SH. “Investigation for software power consumption of code refactoring techniques,” in SEKE, 2014.


