# Investigating the Use of a Hybrid Search Strategy for Systematic Reviews

Erica Mourão[1], Marcos Kalinowski[2], Leonardo Murta[1], Emilia Mendes[3], Claes Wohlin[3]

[1]Fluminense Federal University
Niterói, Brazil
emourao@ic.uff.br;
leomurta@ic.uff.br

[2]Pontifical Catholic University of Rio de Janeiro (PUC-Rio)
Rio de Janeiro, Brazil
kalinowski@acm.org

[3]Blekinge Institute of Technology
Karlskrona, Sweden
emilia.mendes@bth.se;
claes.wohlin@bth.se

***Abstract*—[Background] Systematic Literature Reviews (SLRs) are one of the important pillars when employing an evidence-based paradigm in Software Engineering. To date most SLRs have been conducted using a search strategy involving several digital libraries. However, significant issues have been reported for digital libraries and applying such search strategy requires substantial effort. On the other hand, snowballing has recently arisen as a potentially more efficient alternative or complementary solution. Nevertheless, it requires a relevant seed set of papers. [Aims] This paper proposes and evaluates a hybrid search strategy combining searching in a specific digital library (Scopus) with backward and forward snowballing. [Method] The proposed hybrid strategy was applied to two previously published SLRs that adopted database searches. We investigate whether it is able to retrieve the same included papers with lower effort in terms of the number of analysed papers. The two selected SLRs relate respectively to elicitation techniques (not confined to Software Engineering (SE)) and to a specific SE topic on cost estimation. [Results] Our results provide preliminary support for the proposed hybrid search strategy as being suitable for SLRs investigating a specific research topic within the SE domain. Furthermore, it helps overcoming existing issues with using digital libraries in SE. [Conclusions] The hybrid search strategy provides competitive results, similar to using several digital libraries. However, further investigation is needed to evaluate the hybrid search strategy.**

***Keywords*—*systematic review; search strategy; hybrid strategy; snowballing.***

## I. INTRODUCTION

Evidence-based software engineering is a foundation for empirical software engineering researchers and practitioners. In this context, Systematic Literature Reviews (SLRs) have been used to identify, evaluate and interpret research relevant in several Software Engineering (SE) topics [11]. Guidelines for conducting SLRs in the SE domain have been provided [11], which include the use of database searches to identify studies.

Indeed, the most popular strategy to date in SE SLRs involves employing string-based searches in several digital libraries. Applying this strategy is challenging and involves significant effort. Challenges include the selection of the digital libraries and specific search string adjustments for conducting advanced searches within those libraries. Available digital libraries (e.g., ACM DL, IEEExplore, Scopus, Science Direct and Web of Science) are not designed to support SLRs [3]. They contain significant content overlaps, provide different interfaces, and present search limitations (e.g. on logical operators and number of terms). Moreover, there are usually difficulties in identifying precise terms and synonyms for the search string [16]. As an alternative solution, both Webster and Watson [15] and Wohlin [16] recommend approaches based on snowballing for identifying relevant research on a given topic, mitigating the aforementioned problems. However, the quality of the results obtained by snowballing is most likely highly dependent on the seed set of papers.

The goal of this paper is to propose and carry out a preliminary assessment of a hybrid search strategy that combines both database search and snowballing. This search strategy comprises following the guidelines for preparing a search string based on the proposed research questions [11], conducting a database search in a specific digital library (e.g., Scopus), applying the inclusion and exclusion criteria on the retrieved studies, and then using the selected studies as a seed set for Backward (BS) and Forward Snowballing (FS). This should enable complementing the selected studies with others potentially indexed in different sources. If this strategy enables retrieving relevant studies with a similar recall compared to database searches it would both allow overcoming the challenges faced when using several different digital libraries and also enable identifying a good seed set for BS and FS in a structured way.

For evaluation purposes, we applied the hybrid strategy to two previously conducted SLRs, which employed database searches, to assess whether it would be able to retrieve the same set of included papers with lower effort. We selected two SLRs: one about elicitation techniques in different fields (not confined to SE) and another relating to a specific SE topic on cost estimation.

The remainder of this paper is organized as follows. In Section II, related work is presented. In Section III, we describe the hybrid strategy. The research method and results are respectively presented in Section IV and Section V, and discussed in Section VI. Section VII concludes the paper.

## II. RELATED WORK

There are still few studies evaluating SLR search strategies in SE. Jalali and Wohlin [8] conducted a study to compare the use of database search with snowballing in an SLR investigating agile practices in global software engineering. They concluded that there were no major differences between results. Thereafter, Wohlin [16] proposed guidelines for conducting

snowballing in SLRs via defining and assessing a snowballing procedure through the replication of a previously published SLR where database searches were used. The results suggested that using snowballing, as a search strategy might be a good alternative to using database searches; however results are also dependent on selecting a seed set carefully.

Badampudi *et al.* [1] investigated and compared snowballing with database searches and evaluated the efficiency and reliability of the snowballing approach. In this study, different researchers conducted the database and snowballing strategies independently. They also concluded that the effectiveness of snowballing is comparable to that of database searches.

Wohlin [17] and Felizardo et al. [7] conducted studies investigating snowballing as a search strategy for updating SLRs, i.e., conducting second-generation studies. In their studies, FS was used to replicate a second-generation study previously conducted using database searches. The original SLR and its included studies were used as seed set. The general conclusions indicate that FS was able to find all relevant papers to update the SLR. However, these studies had a different focus and concerned updating previously conducted SLRs, in which previously included studies are already available, making it easier to define a representative seed set.

While some interesting results have been obtained, a hybrid approach has not been formalized yet and further investigations are still needed to better understand under which conditions a certain search strategy should be employed.

## III. THE HYBRID SEARCH STRATEGY

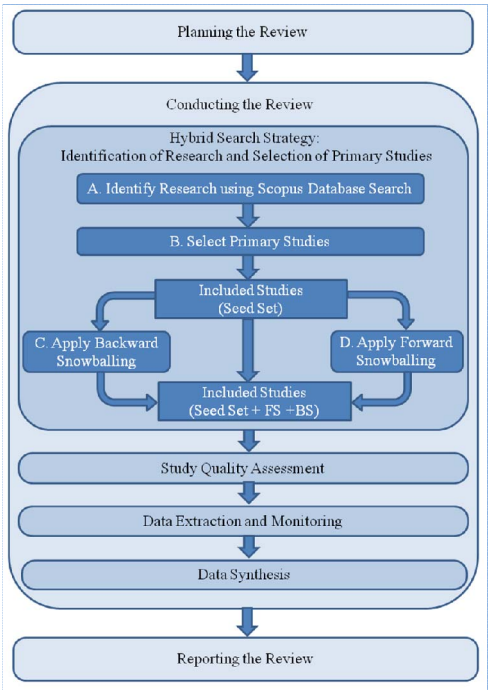
According to Kitchenham et al. [11] the aim of an SLR is to find and summarize all relevant studies addressing one or more research questions using an unbiased search strategy, while Wohlin et al. argue that it is more a matter of how good sample it is possible to identify [19]. Thus, SLRs should be performed in a transparent and replicable way. The SLR process involves several activities and can be summarised in three main phases: *planning*, *conducting the review*, and *reporting the review* [11].

Our proposed hybrid search strategy is to be applied during the *conducting the review* phase, and is shown in Fig. 1. More specifically, it addresses the first two activities that happen within this phase, as per the SLR process in [11]: *identification of research* and *selection of primary studies*. In summary, the strategy comprises identifying and selecting studies by combining searching in only one specific digital library (Scopus) with BS and FS. A detailed description of the hybrid strategy activities is provided in the following subsections. The iterations suggested in [16] are not included in Fig. 1 to avoid a too complex figure.

### A. Identify Research using Scopus Database Search

The first step of the hybrid strategy is to define the search string to identify relevant studies using the chosen digital library. This should be done similarly to traditional SLRs, e.g., based on the PICO (population, intervention, comparison and outcome) strategy [13]. Given that snowballing should be done on the identified papers, it is less critical to identify all possible synonyms as if the database search is the sole driver for identifying relevant papers. Scopus was chosen as the digital library because it covers several disciplines and is defined as the largest abstract and citation database of peer-reviewed literature [6]. Moreover, it presents a clear and consistent search interface.

Fig. 1 Hybrid Search Strategy Process




### B. Select Primary Studies

Once the search has been performed, selection criteria are applied to identify relevant primaries studies [11], hence resulting in a set of candidate studies. The seed set should only include papers that eventually will be included, and hence the quality of the papers needs to be investigated before being used within the seed set [16]. Within the hybrid strategy these studies represent the seed set for applying BS and FS, which are detailed the next.

### C. Apply Backward Snowballing

BS is conducted on the set of studies obtained from activity B and is executed in several steps. The first step is to carefully review the reference list of each paper searching for new studies. New studies should only be included in the SLR if they pass the selection criteria. As described by Wohlin [16], the process continues with new iterations on the papers selected during the previous iteration and ends if there are no new papers to be selected. When conducting BS, a minor deviation from the guidelines [16] was made by not applying FS on the papers found in BS. The deviation disfavours snowballing, and hence it makes the evaluation somewhat conservative.

### D. Apply Forward Snowballing

FS is also conducted from the obtained seed set and divided in iterations. It comprises searching for new studies that cite the studies contained in the seed set [7]. As done by Wohlin [17], we suggest using Google Scholar. Similarly to BS, new studies should only be included in the SLR if they pass the

selection criteria and the process continues with new iterations until there are no new papers to be selected. A similar deviation from the guidelines was done when conducting FS, i.e. BS was not conducted on the papers found in FS.

The result from applying the hybrid strategy is a set of papers that we hypothesize to include the same studies that would have been retrieved by means of database searches using different databases, and, potentially with more effort.

## IV. RESEARCH METHOD

To assess the proposed hybrid approach, we have replicated two previously conducted SLRs in SE that used database searches to select primary studies. Note that the comparison was based on the set of studies included in the previously published SLRs, checking if they were also retrieved by the hybrid strategy. The two SLRs relate to different research topics: one focuses on elicitation techniques in a range of areas, not only SE (this study is called SR1 henceforth), and the other SLR (SLR2 henceforth) focuses on software cost estimation.

### A. Study Goal and Research Questions

The main goal of this research is to assess whether the hybrid search strategy presents preliminary competitive results for conducting SLRs (in particular in SE), when compared to databases searches. Such assessment considers both the efficiency (precision [4]) and the completeness (relative recall [4]) of the hybrid approach when compared to database search. This goal is addressed via the following research questions:

*RQ1: Does the hybrid strategy involve analysing fewer papers when compared to the published SLRs that conducted database searches?* To address this research question, we measure precision as being the percentage of the analysed papers that were included in the SLR. Note that precision is an appropriate metric for representing efficiency since it degrades in face of noisy results. For example, should all analysed papers be included, precision would be 100%, showing high efficiency. On the other hand, should most analysed papers be discarded, precision would tend to zero, indicating low efficiency.

*RQ2: Does the hybrid strategy retrieve all the papers that were included in the published SLRs that conducted database searches?* To address this research question, we measure the relative recall as being the percentage of included papers of the published SLR that were retrieved by the hybrid research strategy.

### B. Instrumentation and Procedure

We strategically selected two SLRs for applying de hybrid approach, one with studies surpassing the SE domain and another one with studies limited to the SE domain. This would allow us to observe if there would be differences for the SE domain. Moreover, both were considered interesting candidates given that they were published in high-quality venues, were conducted using database search strategies involving several digital libraries, and documented the search string and intermediate results.

SLR1, by Dieste and Juristo [5], includes empirical studies on elicitation techniques aiming at providing knowledge about elicitation techniques in general, rather than only in SE. It was published in 2011, with a search period including all studies published up to March 2006 (inclusive). The search strategy involved designing and running a search string on ACM DL, IEEE Xplore, and Scopus (according to the authors, these searches were complemented with some opportunistic searches conducted in Google Scholar). SLR1 identified 26 research papers that met its inclusion criteria. The list of included papers of SLR1 is available online[1].

SLR2, by Mendes et al. [12], includes studies comparing the effort prediction accuracy between cross- and within-company prediction models. SLR2 updates a previously conducted (SLR) comprising a search period including studies published up to 2006 [9] with additional studies published between 2006 to November 2013 (inclusive) [12]. The update used the exact same search strategy, which relied on designing and running a search string on the following digital libraries: ACM DL, EI Compendex, IEEE Xplore, Inspec, Science Direct, Scopus, and Web of Science. In total, 25 research papers (published before November 2013) that met the SLR inclusion criteria are retrieved by this search strategy. The list of these 25 papers is also available online[1].

When applying the hybrid approach, note that during snowballing, we limited our effort to identifying, within the new candidate studies, the ones that were included in the original SLRs. Thus, as mentioned before, we focused on effort and capability of reproducing equivalent results of the original SLRs, i.e., our research questions do not investigate if the hybrid strategy would retrieve more/other papers to be included than the published SLR. At the end, we calculated the precision and relative recall to address the posed research questions.

## V. RESULTS

This section presents the results of our case studies, addressing research questions *RQ1* and *RQ2*.

### A. Case Study I: SLR on Requirements Elicitation Techniques

#### 1) Applying the Hybrid Strategy

We replicated the original database search on Scopus to create the seed set for snowballing. During this step, 15 of the 26 studies included in SLR1 were retrieved (out of 737 results). Next, following the hybrid strategy, we applied BS on these 15 studies (see Table I). In the first iteration four more included studies were identified within the 425 references. In the second iteration, we found no more studies within the 125 references of the 4 studies identified in the first iteration. Thus, a total of 19 studies were found using Scopus + BS (73% of the studies).

Thereafter, we applied FS on the 15 studies of the seed set. In the first iteration, we found two more studies included in SLR1 (see Table II) within 459 citing papers. In the second iteration, we found one additional study within the 36 papers citing the two studies identified in the first iteration. In the third iteration we analysed two citing papers and found no more included studies. Hence, a total of 18 studies were found using Scopus + FS (69% of studies).

[1] http://www.wohlin.eu/esem17

In summary, the hybrid search strategy allowed identifying a total of 21 out of the 26 studies included in SLR1. The precision of Scopus + BS + FS was 21/1804 and the relative recall 21/26 (81%). Further explanations concerning the missed studies are presented in the discussion in Section V.

TABLE I. CASE STUDY 1: BS – ITERATIONS 1 AND 2

| Studies | | Backward Snowballing (BS) - iteration 1 | | | | | | | | | | | | | | | BS - iteration 2 | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Study found in Scopus (Seed Set) | | | | | | | | | | | | | | | Study found in BS - iteration 1 | | | |
| | | S1 | S3 | S4 | S5 | S7 | S9 | S11 | S13 | S14 | S15 | S16 | S18 | S19 | S21 | S24 | S2 | S6 | S10 | S20 |
| Study not found in Scopus | S2 | | | | X | | | | | | | X | | | | | | | | |
| | S6 | | | | | X | | | | | X | | | | X | | | | | |
| | S8 | | | | | | | | | | | | | | | | | | | |
| | S10 | | | | X | | | | | | | | | | | | | | | |
| | S17 | | | | | | | | | | | | | | | | | | | |
| | S20 | | | | X | | | | | | | | | | | | | | | |
| | S22 | | | | | | | | | | | | | | | | | | | |
| | S23 | | | | | | | | | | | | | | | | | | | |
| | S25 | | | | | | | | | | | | | | | | | | | |
| | S26 | | | | | | | | | | | | | | | | | | | |

TABLE II. CASE STUDY 1: FS – ITERATIONS 1, 2, AND 3

| Studies | | Forward Snowballing (FS) - iteration 1 | | | | | | | | | | | | | | | FS - iteration 2 | | FS - iteration 3 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Study found in Scopus (Seed Set) | | | | | | | | | | | | | | | Study found in FS - iteration 1 | | Study found in FS - iteration 2 | |
| | | S1 | S3 | S4 | S5 | S7 | S9 | S11 | S13 | S14 | S15 | S16 | S18 | S19 | S21 | S24 | S20 | S22 | S22 | S26 |
| Study not found in Scopus | S2 | | | | | | | | | | | | | | | | | | | |
| | S6 | | | | | | | | | | | | | | | | | | | |
| | S8 | | | | | | | | | | | | | | | | | | | |
| | S10 | | | | | | | | | | | | | | | | | | | |
| | S17 | | | | | | | | | | | | | | | | | | | |
| | S20 | | | | | | | | | | | | X | | | | | | | |
| | S22 | | | | X | | | | | | | | X | | | | X | | | |
| | S23 | | | | | | | | | | | | | | | | | | | |
| | S25 | | | | | | | | | | | | | | | | | | | |
| | S26 | | | | | | | | | | | | | | | | X | | | |

*2) Addressing the Research questions*

*RQ1: Does the hybrid strategy involve analysing fewer papers when compared to the published SLRs that conducted database searches?*

The hybrid search strategy identified a total of 21 studies out of 26, after analysing 1804 studies. On the other hand, using the database search approach in SLR1, 4062 papers were retrieved from the digital libraries. It is noteworthy that the authors had to significantly adjust the search string to handle differences of the digital libraries. Moreover, they performed additional opportunistic and unspecified searches using Google Scholar. The paper does neither document which (if any) of the included papers were identified from those additional searches, nor the number of analysed papers in these searches. Consequently, 4062 is an underestimate of the actual number of papers to be analysed using the database search strategy.

Thus, with regard to SLR1, the answer to RQ1 was *yes*, it involved analysing fewer papers and provided a higher precision (21/1804 against 26/4062).

*RQ2: Does the hybrid strategy retrieve all the papers that were included in the published SLRs that conducted database searches?*

The answer to RQ2 was *no*, given that the hybrid strategy retrieved 21 out of the 26 papers included in SLR1. Thus, relevant primary studies might have been lost during the process. We believe that this could be due to the large study population, including papers from different areas, which do not tend to cite each other or the extra search conducted by the authors. Independently, it points to the importance of the seed set in terms of coverage of different categories of papers [1].

## B. Case Study II: SLR on Cross- vs Within-Company Cost Estimation

*1) Applying the Hybrid Strategy*

We replicated the original database search employed in SLR2 on Scopus to create the seed set for snowballing; verifying which studies of the list of 25 included studies identified by SLR2's search strategy would also be retrieved by the hybrid strategy. During this step, 17 of the 25 included studies were retrieved (out of 603 results). Following the hybrid strategy, we applied BS on these 17 studies (see Table III). In the first iteration we found six additional studies within the 422 references listed in the papers of the seed set. In the second iteration, 148 references were analysed and no additional studies were found. Thus, a total of 23 studies were found using Scopus + BS (92% of the studies).

Thereafter, we applied FS on the 17 studies of the seed set. In first iteration we found all eight additional studies (see Table IV) analysing 1089 citing papers. Hence, all studies were found using Scopus + FS (100% of the studies).

In summary, for this case study the hybrid search strategy identified all relevant studies. The precision of Scopus + BS + FS was 25/2262 and the recall 25/25 (100%).

TABLE III. CASE STUDY 2: BS – ITERATIONS 1 AND 2

| Studies | | Backward Snowballing (BS) - iteration 1 | | | | | | | | | | | | | | | | | BS - iteration 2 | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Study found in Scopus (Seed Set) | | | | | | | | | | | | | | | | | Study found in BS - iteration 1 | | | | | |
| | | S1 | S2 | S4 | S5 | S6 | S7 | S9 | S10 | S11 | S12 | S14 | S16 | S18 | S19 | S20 | S21 | S23 | S3 | S8 | S13 | S15 | S17 | S22 |
| Study not found in Scopus | S3 | | | | X | | | | | X | | | | | | | | | | | | | | |
| | S8 | | | | | | | X | X | X | X | X | X | | | X | | | | | | | | |
| | S13 | | | | | | | | | | | | X | | | | X | | | | | | | |
| | S15 | | | | | | | | | | | | | | | X | | | | | | | | |
| | S17 | | | | | | | | | | | | | | | | X | | | | | | | |
| | S22 | | | | | | | | | | | | X | | X | | | | | | | | | |
| | S24 | | | | | | | | | | | | | | | | | | | | | | | |
| | S25 | | | | | | | | | | | | | | | | | | | | | | | |

TABLE IV. CASE STUDY 2: FS – ITERATION 1

| Studies | | Forward Snowballing (BS) - iteration 1 | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Study found in Scopus (Seed Set) | | | | | | | | | | | | | | | | |
| | | S1 | S2 | S4 | S5 | S6 | S7 | S9 | S10 | S11 | S12 | S14 | S16 | S18 | S19 | S20 | S21 | S23 |
| Study not found in Scopus | S3 | | X | | | | | | | | | | | | | | | |
| | S8 | | X | X | X | X | | | | | | | | | | | | |
| | S13 | | X | X | X | X | X | X | X | X | | | | | | | | |
| | S15 | X | X | X | X | X | X | X | X | X | X | | | | | | | |
| | S17 | | X | X | X | X | X | X | X | X | | X | X | | | | | |
| | S22 | X | X | X | X | X | | X | X | | | | | | | | | |
| | S24 | | | | | | | | | | | | | X | | | X | X |
| | S25 | | | | | | | X | | | | | | | | | | |

*2) Addressing the Research questions*

*RQ1: Does the hybrid strategy involve analysing fewer papers when compared to the published SLRs that conducted database searches?*

The hybrid search strategy identified all 25 studies after analysing 2262 studies. On the other hand, using the database search approach (considering the 772 of the original SLR [9] and 980 of the update [12]), 1752 papers were analysed. However, the original SLR [9] also conducted manual searches for which the exact numbers are not available. Moreover, the search string had to be adapted significantly for several digital libraries (in ACM DL, for example, it would have returned 386442 papers). Nevertheless, considering the adjustments

made to the search string, the answer to RQ1 in this case study is *no,* it involved analysing more papers and had a lower precision (25/2262 against 25/1752). It is noteworthy, however, that the hybrid approach involves a simpler search process and did not require handling awkward adjustments to adapt to differences in the digital libraries.

*RQ2: Does the hybrid strategy retrieve all the papers that were included in the published SLRs that conducted database searches?*

The answer for RQ2 in this case study was *yes*, given that the hybrid strategy retrieved all 25 papers included in SLR2. We believe that this behaviour could be related to having a narrower study population, where relevant papers tend to cite each other more consistently.

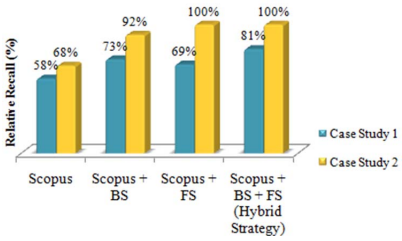
## VI. DISCUSSION

### A. Reliability of the Hybrid Search Strategy

Reliability of a search strategy mainly regards relevant studies that were missed. Thus, we use the relative recall (percentage of previously known relevant studies that were identified) as a basis for this discussion.

The relative recall of the hybrid strategy's search steps for both case studies is shown in Fig. 2. As mentioned before, we believe that the higher results for case study 2 could be due to having a narrower study population, where relevant papers tend to cite each other more consistently (citation graphs for both SLRs are available online[1]). Thus, a preliminary interpretation is that the hybrid strategy may be more reliable for more narrow study populations and not recommended for large study populations covering a broader range of domains. It may also be argued that it illustrates the need to have a good seed set, for example including different categories of papers being representative of the area studied [1].

Fig 2. Relative recall of the hybrid search strategy in case study 1 and 2.



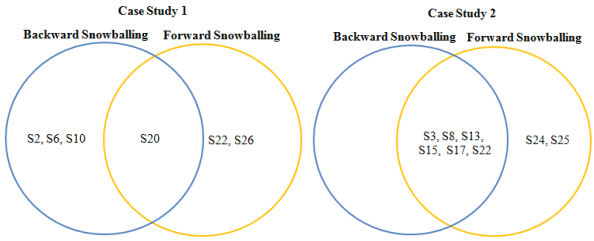

To further understand the reliability, we analysed the five studies of SLR1 that were not retrieved by the hybrid strategy (S8, S12, S17, S23, and S25). S8 is a short paper dating from 1989 from outside the area of SE (knowledge engineering) that is not indexed in Scopus. S12 is a paper on SE dating from 1990 that has received relatively few citations. The paper is indexed in Scopus, but was not retrieved by Scopus using the search string designed by the SLR authors. S17 is a paper from the area of Psychology, which is indexed in Scopus but was also not retrieved by Scopus using the search string. Finally, papers S23 and S25 are from the area of marketing. S23 is not indexed by Scopus but well cited and S25 is grey literature (a master dissertation). Thus, three papers not found using the hybrid approach are from other disciplines (S8, S17 and S23), one paper is from SE (S12) but with relatively few citations and finally one paper is from the grey literature (S25).

In summary, it seems like four papers included in SLR1 (S12, S17 and S23 indexed in Scopus, but not found using the search string when applied in Scopus; and S25 being grey literature) that were not found by the hybrid search strategy may indeed have been found in the additional opportunistic database searches conducted by the authors of SLR1. Those opportunistic searches resulted in 11 additional papers being selected for filtering (we do not have the information on which of those ones, if any, were included, although four candidates are identified above). Given this reasoning, the precision and recall of SLR1's database search strategy may be discussed.

It is noteworthy that the search strings of SLR1 and SLR2 had to be adapted for the different digital libraries. In particular, the ACM DL required significant adaptations in both cases until retrieving a manageable amount of results. This illustrates some of the difficulties in handling several digital libraries with sometimes limited and/or inconsistent search mechanisms.

It is also possible to observe in Fig. 2 that BS and FS do not replace each other and may both contribute to the recall of relevant papers. This makes sense, given that BS may help identifying older references while FS favours finding newer ones. Since the hybrid strategy involves selecting a seed set from only one digital library in which older and newer references that are not indexed might still be missing, ideally BS and FS should both be applied according to the guidelines [16] when considering the hybrid strategy. The Venn diagram shown in Fig. 3 illustrates the studies found using BS and FS in both case studies, and reinforces this argument.

Fig 3. Case Study 1 e 2: Venn Diagram – BS and FS.




In summary, given the issues faced when using several digital libraries (e.g., library selection, different interfaces, inconsistent search mechanisms, search-string limitations, required adjustments, and relying on effective identification of synonyms of terms) [16] and our obtained results on relative recall and precision we believe that the hybrid strategy potentially represents a reasonable alternative when conducting an SLR on a specific topic within the SE domain. However, additional evaluations are required before drawing further conclusions. Considerations on the limitations of this research follow.

### B. Threats to Validity

As in all studies, this research has certain limitations. The main limitation is that our analyses were based on evaluating the hybrid strategy on only two SLRs, so further evaluation is clearly needed. However, the results have provided some

valuable insight into how much the hybrid approach can be used as a search strategy for SLRs. Thus, as a preliminary evaluation, we believe our case studies met their goal. Threats to validity are described hereafter.

*1) Internal validity:* The hybrid approach was defined based on the SLR process described in the guidelines for conducting SLRs [11] and for applying it we followed the guidelines for conducting snowballing [16] with a minor deviation. The first author applied the approach and each step was reviewed by at least one of the other authors.

*2) Construct validity:* In our study we only verified which of the included studies were also retrieved by the hybrid strategy, and hence the authors of the original SLRs applied all inclusion and exclusion criteria. Moreover, the use of relative recall may miscount true positive studies obtained by the hybrid approach that were not included in the original SLR. Consequently, the precision and the relative recall of the hybrid approach are lower bounds.

*3) Reliability:* We evaluated the hybrid strategy using Scopus and Google Scholar for FS. The evaluation process was straightforward, clearly described and replicable.

*4) External validity:* The study of applying the hybrid strategy was conducted on two peer reviewed SLRs. However, the study findings are not generalizable and replications on other SLR to reinforce our findings are required.

## VII. CONCLUSION

The goal of this paper was to formalise a hybrid search strategy and evaluate the feasibility of applying it to support finding evidence when conducting SLRs. The hybrid strategy involves combining searching in a specific digital library (Scopus) with backward and forward snowballing. Thereby, mitigating some of the challenges with managing different databases and identifying all synonyms in database searches, and the challenge of finding a suitable seed set for snowballing. Jalali and Wohlin [8] highlight some of the challenges with synonyms when comparing database searches vs. snowballing, while Badampudi et al [1] discusses some of the challenges with an appropriate seed set when using snowballing.

We evaluated the hybrid strategy for two SLRs, analysing precision and relative recall compared to conducting searches in several digital libraries. Our results provide preliminary indications that it may be a suitable search strategy under certain circumstances. However, it is too early, given that only two SLRs have been evaluated, to draw stronger conclusions.

Nevertheless, our results also indicated that for SLRs with a larger study population (for example, surpassing the SE domain) the hybrid strategy might result in lower relative recall, implying the risk of missing relevant studies. However, SLR1 is in retrospect not optimal for comparison given the additional searches done by the authors of SLR1.

Additional evaluations are required to draw further conclusions and are part of future work, which also comprises comparing the hybrid approach to BS and FS applied to differently obtained seed sets, as well as evaluating if the hybrid approach helps mitigating some of the challenges with database searches respectively snowballing.

## REFERENCES


[1] D. Badampudi, C. Wohlin and K. Petersen, “Experiences from using snowballing and database searches in systematic literature studies”, Proceedings Evaluation and Assessment in Software Engineering, 2015.

[2] V.R. Basili and H.D. Rombach, “The TAME project: towards improvement-oriented software environments,” IEEE Transactions on Software Engineering, 14(6), pp. 758–773, 1988.

[3] P. Brereton, B. Kitchenham, D. Budgen, M. Turner, and M. Khalil, “Lessons from applying the systematic literature review process within the software engineering domain”, Journal of Systems and Software, 80 (4), pp. 571–583, 2007.

[4] K. Dickersin, R. Scherer, and C. Lefebvre, “Systematic reviews: Identifying relevant studies for systematic reviews”, British Medical Journal, 309 (6964), pp. 1286–1291, 1994.

[5] O. Dieste and N. Juristo, “Systematic Review and Aggregation of Empirical Studies on Elicitation Techniques”, In IEEE Transactions on Software Engineering (TSE), vol. 37, no.2, march/april 2011.

[6] Elsevier, “Scopus”, https://www.elsevier.com/solutions/scopus. Accessed in 04/2017.

[7] K. R. Felizardo, E. Mendes, M. Kalinowski, E. F. Souza, and N. L. Vijaykumar, “Using forward snowballing to update systematic reviews in software engineering”, International Symposium on Empirical Software Engineering and Measurement (ESEM’16), pp. 53:1–53:6, 2016.

[8] S. Jalali and C. Wohlin, "Systematic Literature Studies: Database Searches vs. Backward Snowballing", International Conference on Empirical Software Engineering and Measurement, ESEM’12, September 19-20, 2012.

[9] B. A. Kitchenham, E. Mendes, and G. H. Travassos, “Cross versus within-company cost estimation studies: A systematic review”. IEEE Transactions on Software Engineering, 33 (5), pp. 316–329, 2007.

[10] B. A. Kitchenham, T. Dyba, and M. Jorgensen., “Evidence-based software engineering”, Proceedings. 26th International Conference on Software Engineering, pp. 273-281, 2004.

[11] B. A. Kitchenham and S. Charters, “Guidelines for performing systematic literature reviews in software engineering”, Technical Report EBSE 2007–001, Keele University and Durham University Joint Report, 2007.

[12] E. Mendes, M. Kalinowski, F. Ferrucci and F. Sarro, “Cross-vs. Within-Company Cost Estimation Studies Revisited: An Extended Systematic Review”, In 18th Int. Conf. on Evaluation and Assessment in Software Engineering, (EASE’14), 2014.

[13] M. Pai, M. McCulloch, J.D. Gorman, N. Pai, W. Enanoria, G. Kennedy, P. Tharyan, and J. M. Colford, Jr., “Systematic reviews and meta-analyses: An illustrated stepby-step guide”, The National Medical Journal of India, 17(2), 86-95, 2004.

[14] P. Runeson, M. Höst, and A. Rainer, “Case Study Research in Software Engineering: Guidelines and Examples”, J. Wiley, Ed., 2012.

[15] J. Webster and R. Watson, “Analyzing the past to prepare for the future: Writing a literature review”, MIS Quarterly, vol. 26, nº. 2, pp. 13–23, 2002.

[16] C. Wohlin, “Guidelines for Snowballing in Systematic Literature Studies and a Replication in Software Engineering”, International Conference on Evaluation and Assessment in Software Engineering, 2014.

[17] C. Wohlin, “Second-generation Systematic Literature Studies Using Snowballing”. International Conference on Evaluation and Assessment in Software Engineering, pp. 15:1–15:6, 2016.

[18] C. Wohlin, “Writing for synthesis of evidence in empirical software engineering”, In 8th Int. Symp. on Empirical Software Engineering and Measurement (ESEM’ 14), 2014.

[19] C. Wohlin, P. Runeson, P. A. da Mota Silveira Neto, E. Engström, I. do Carmo Machado and E. S. de Almeida, "On the Reliability of Mapping Studies in Software Engineering", Journal of Systems and Software, Vol. 86, No. 10, pp. 2594-2610, 2013.

[20] X. Zhou, Y. Jin, H.Zhang, S. Li and X. Huang, “A Map of Threats to Validity of Systematic Literature Reviews in Software Engineering”, 23rd Asia-Pacific Software Engineering Conference 2016.