A Systematic Mapping Study on Formal Methods Applied to Crosscutting Concerns Mining

Rafael S. Durelli and Vinícius H. S. Durelli

Computer Systems Department
University of São Paulo
São Carlos – Brazil
{rdurelli, durelli}@icmc.usp.br

Abstract. Background: Crosscutting concerns consist in software system features having the implementation spread across modules as tangled and scattered code. Developers need of up-to-date knowledge about crosscutting concerns currently implemented in their systems and about the location of these concerns throughout the underlying code. Hence, within the academic community there are a bunch of techniques to mine such scattered concerns. To the best of our knowledge, there is no survey reporting on what methods have been employed in this area.
Objectives: To conduct a systematic mapping study to ascertain what formal methods have been used to assist the mining of crosscutting concerns.
Research Method: We have carried out a systematic mapping study of the literature based upon searching of major electronic databases.
Results: As a result, 10 primary studies have been selected and classified by their categories and data of publication. From analyzing the results of our mapping study, we found out that formal methods have scarcely been used to assist the identification of crosscutting concerns.
Conclusions: According to our results, there is still much research to be done on applying formal methods to support crosscutting mining. Therefore, this mapping study may be seen as an initial step towards identifying research gaps in the area and thus perspectives for future research.

Keywords: Systematic Mapping, Formal Method, Crosscutting Concerns

1 Introduction

A concern is commonly defined as anything that stakeholders regard as a conceptual unit [1]. Concerns range from development oriented tracing, and more general purpose caching, to domain-specific business rules. Developers and architects continuously need of up-to-date knowledge about concerns currently implemented in their systems as well as their location throughout the code. For

example, during maintenance and reengineering, developers need to locate specific concerns in the source code. Bug fixes must be propagated to the whole implementation of a concern [2], and possibly to other concerns with which the concern interacts. Architects need to map the currently implemented concerns to the reference architecture to verify architecture conformance. As such, crosscutting concern mining is indispensable for software maintenance, reverse engineering, reengineering and even for re-documentation. Despite their importance, as far as we know the literature, there are no surveys describing what techniques have been used to identify crosscutting concerns. Therefore, as a first step towards filling this gap, the main objective of this paper is to present a systematic mapping that attempts to identify whether formal methods have been used to aid the identification of crosscutting concerns.

2 The Mapping Process

The process applied to conduct the mapping study herein described is detailed by [3]. According to them, its essential steps are: *(i)* definition of research questions, *(ii)* conducting the search for relevant primary studies, *(iii)* screening of papers, *(iv)* keywording of abstracts, and *(v)* data extraction and mapping. Each step produces an intermediate outcome, the concluding result being the mapping study as shown in Figure 1. Following sections present details on how each step was carried out.

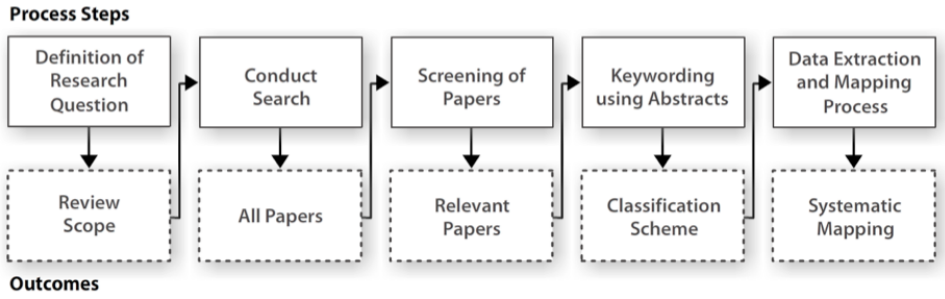

Fig. 1. Systematic mapping process. This Figure was adapted from [3]

2.1 Definition of Research Questions

Research questions must embody the mapping study purpose. Moreover, research questions reflect the general scope of the mapping study. The scope is comprised of population (i.e., population group observed by the intervention), intervention (i.e., what is going to be observed in the context of the planned mapping study), and outcomes of relevance (i.e., the results of the intervention). Accordingly,

during the conduction of this step, it was also necessary to establish the scope of the mapping study:

- **Population**: published scientific literature reporting on formal methods applied to aspects mining or/and crosscutting mining.
- **Intervention**: published scientific literature concerned with introducing formal methods related to aspects mining or/and crosscutting concerns mining. Furthermore, we also aim at determining whether there is a formal method that is applied in this domain and if it is used within academic circles.
- **Outcome of relevance**: an overview of the studies that have been conducted in the domain of formal methods, emphasizing primary studies that report on how these methods can be used to aid the identification of crosscutting concerns. From observing such an aggregated data set we also intend to provide insight into the frequencies of publication over time to inspect trends.

Hence, given that we set out to determine if formal methods are being applied to identify crosscutting concerns, our research question (RQ) reflects this purpose as follows:

- $\mathbf{RQ}_1$: Have formal methods being used to support the mining of crosscutting concerns within the academic community?

2.2 Search for Primary Studies

Herein the search for primary studies basically involves defining both the search string and electronic databases to be used. The string we used for searching is composed of a combination of the following keywords and acronyms: ***concern mining, aspect mining, mining technique, crosscutting concerns and formal method***. Figure 2 depicts the search string devised from combining such keywords.

*(("aspect mining") **OR** ("concern mining") **OR** ("CC mining") **OR** ("mining technique") **OR** ("crosscutting concerns") **OR** ("crosscutting concerns")) **AND** (("formal method") **OR** ("formal-method") **OR** ("formal specification") **OR** ("formal verification"))*

Fig. 2. Search string.

Afterwards, we used such search string on the following electronic databases: ACM Digital Library[1], IEEE Xplore[2], ScienceDirect[3] and Springer Lecture Notes

[1] `http://portal.acm.org`
[2] `http://ieeexplore.ieee.org`
[3] `http://www.sciencedirect.com`

in Computer Science (LNCS)[4]. These databases have search engines that enable identifying for occurrences of the terms defined both in the title and abstract. Furthermore, no limits were placed on date of publication in order not to restrict the mapping study scope. Aimed at keeping track of the selected papers, we used JabRef[5], an open source system for bibliography reference management.

2.3 Screening of Papers for Inclusion and Exclusion Criteria

In order to determine which primary studies are relevant to answer our research question, we applied a set of inclusion and exclusion criteria to each retrieved study.

Inclusion criteria (IC) devised and applied are:

1. if several paper reported identical studies, only the most recent was selected;
2. papers describing more than one study had each study individually evaluated;
3. the primary study has to describe at least one technique that uses formal method or model checking or formal verification in the domain of crosscutting mining or aspect mining.

The set of exclusion criteria (EX) are:

1. papers that do not present studies related to formal methods in the domain of crosscutting or aspect mining techniques;
2. papers that do not describe studies related to either formal methods or crosscutting/aspect mining techniques;
3. the primary study is not available in an electronic format;
4. technical reports, documents that are available in the form of either abstracts or presentations (i.e., elements of "grey" literature), and secondary literature reviews (i.e., systematic literature reviews and mapping studies).

It is worth highlighting that we avoided imposing many restrictions on primary study selection since we wanted a broad overview of the research area as a whole. Taking into consideration only certain types of studies would lead to a biased overview and result in an inaccurate mapping study.

Although mapping studies are often carried out based solely on the abstracts, throughout primary study selection these criteria were applied to the following sections of each candidate study: *(i)* title, *(ii)* abstract, *(iii)* introduction, and *(iv)* conclusion. Overall, an initial figure of 11 candidate papers was obtained after applying the inclusion and exclusion criteria based only upon title and abstract. After going over introductions and conclusions, we ended up with a final set of 10 primary studies as shown in Table 1. Nevertheless, it is important to mention that other parts, besides introductions and conclusions, quite often had to be read in order to ascertain whether formal methods were used in the

[4] `http://www.springer.com/lncs`
[5] `http://jabref.sourceforge.net/`

domain of crosscutting/aspect mining. The titles, references, and the inclusion criteria (IC) that was applied for these primary studies are listed in Table 2.

Despite the reduced number of primary studies, we believe that our search retrieved all primary studies that deal with formal methods, formal verification and model-checking within the domain of crosscutting/aspect mining.

Table 1. Papers retrieved from each electronic database, total of candidate studies and the final set.

Electronic Database	Number
ACM Digital Library	93
IEEE Xplore	253
ScienceDirect	121
Springer LNCS	47
Total	514
Candidates	11
Final set	10

2.4 Keywording

As previously mentioned, as far as we know the literature, there is no similar study providing a taxonomy and an overview of research into formal methods and crosscutting/aspect mining. Hence, we applied a *keywording* strategy aimed at devising our own classification scheme and categories for the selected primary studies. Keywording reduces the time spent developing classification schemes and categories. By applying such strategy, initially, abstracts are read for the purpose of finding keywords and concepts that reflect their contribution. Subsequently, these keywords and concepts are combined together to produce a general understanding regarding the nature and contribution of the research. Eventually, the final set of keywords is used to define representative categories. The classification scheme gradually evolves toward its final version as new categories are added, merged, or split up. At the end of the depicted strategy, 3 categories have been obtained: *"**Model-Checking**"*, *"**Formal Verification**"* and *"**Formal Method**"*.

2.5 Data Extraction and Mapping

Each included primary study was assigned to one or more categories that we have devised in the foregoing step. Given these categories, in this section we outline each of the resulting categories as well as typical primary studies of each of them.

– **Model-Checking**: In this category are included primary studies that report on how to use model-checker in the domain of crosscutting/aspect mining.

As evidenced by the included studies there is a lack of formal methods in the domain of aspect oriented (AO) due to the fact that it is not straightforward to implement such methods in this paradigm. Only 3 studies explore how to use model-checker in such domain. In [4] the authors propose a model-checker to verify the correctness of AO programs. Similarly, in [5, **?**] the authors describe theoretical underpinnings for applying model checking to programs written using AO programming languages. Nevertheless, there were no studies reporting on utilizing model-checkers to support crosscutting concerns mining.

- **Formal Verification**: This category comprises studies focusing on introducing functionality for formal verification into crosscutting/aspect mining. For example, in [6] an approach to formally verifying properties of systems composed of multiple crosscutting concerns is presented. This approach models concerns as set of concurrent process and provides a method of composition that mimics the composition operators of existing multiple concern implementation language. In a similar way, in [7] propose an approach that uses the method B to support formal verification in the AO paradigm. Besides formal verification, this category also groups together primary studies concerned with formal specification of crosscutting concerns [8].
- **Formal Method**: This category contains primary studies focusing on providing support for new formal method in the context of crosscutting/aspect, as the study described in [9]. They have implemented a new formal method named AOZCL, an aspect-oriented extension to a formal framework (ZCL) with a built-in software architecture description language.

3 Analysis and Classification

The focus of this section is to present a broad overview of research within formal method considering the context of crosscutting and aspect mining we have acquired after classifying and categorizing primary studies. For this, we analyzed each paper, through the reading of title, keywords and abstract.

During the analysis step, we have identified three categories: *Model-Cheking*, *Formal Verification*, and *Formal Method*. We also used information drawn from this overview to answer our mapping study's research question. The fan plot in Figure 3 (b) depicts the number of primary studies according to their categories. It is worth highlighting that certain primary studies were grouped in more than one category, which affected the frequency count: the sum of the frequencies, i.e., 12, is greater than the total of selected studies presented in Table 1 (10).

As can be seen, the majority of the selected primary studies are published by ACM as shown in Figure 3 (a). The other electronic databases, IEEE, ScienceDirect, and Springer had 2, 2, and 1 selected studies, respectively.

It is fairly evident from observing Figure 3 (c) that the categories formal verification and model checking have not drawn much attention in the last decade. In the same way, formal methods have not been an active research as can be observed from 2001 up to 2006. Nevertheless, our results show that from 2006 up

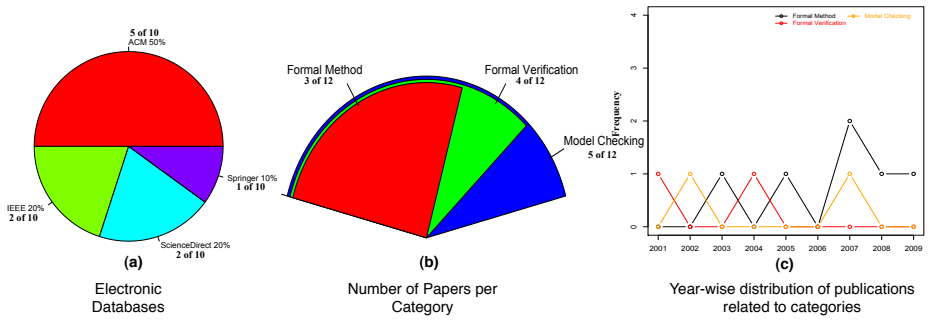

Fig. 3. Distribution of primary studies by electronic database.

to 2009 there has been an increase in the number of publications addressing this subject. In spite of reduced number of primary studies identified herein we argue that the answer to **RQ1** is that *formal methods*, *model checking* and *formal verification* have yet to enter the mainstream. As of now, they have been scarcely used to help identify crosscutting concerns within the academic community.

4 Threats to Validity

The threats to validity in this systematic mapping are the following:

Primary studies selection: Aimed at ensuring an unbiased selection process we defined research questions in advance and devised inclusion and exclusion criteria we believe are detailed enough to provide an assessment of how the final set of primary studies was obtained. However, we cannot rule out threats from a quality assessment perspective, for we simply selected studies without assigning any scores. In addition, we wanted to be as inclusive as possible, thus no limits were placed on date of publication and we avoided imposing many restrictions on primary study selection since we wanted a broad overview of the research area.

Missing important primary studies: The search for primary studies was conducted in several search engines, even though it is rather possible we have missed some primary studies. Nevertheless, this threat was mitigated by selecting search engines which have been regarded as the most relevant scientific sources [10] and therefore prone to contain the majority of the important studies.

Keywording reliability: The authors are software engineering researchers, and one of them is familiar with crosscutting mining, aspect mining, and software product line. However, none of them has theoretical or practical knowledge about formal methods, model-checking, and formal verification, thereby it is possible that we end up introducing some bias during the keywording step. Nevertheless, we believe that our search retrieved all primary studies that deal with formal methods, formal verification, and model-checking within the domain of crosscutting and aspect mining.

Table 2. Selected primary studies

#	Title and Reference	IC
1	**Locating crosscutting concerns in the formal specification of distributed reactive systems**. [11]	IC3
2	**Verifying aspect advice modularly**. [5]	IC3
3	**Foundations of incremental aspect model-checking.** [12]	IC3
4	**Revisiting a Formal Framework for Modeling Aspects in the Design Phase.** [9]	IC3
5	**Aspect-oriented programming with model checking.** [4]	IC3
6	**Using B to Verify the Weaving of Aspects.** [7]	IC3
7	**An Approach for Modeling and Analyzing Crosscutting Concerns.** [8]	IC3
8	**Supporting Formal Verification of Crosscutting Concerns.** [6]	IC3
9	**Coordinating Aspects and Objects.** [13]	IC3
10	**Using Aspects for Enforcing Formal Architectural Invariants.** [14]	IC3

5 Concluding Remarks

The main contribution of this paper is an overview of the formal methods that have been used to help the mining of crosscutting concerns. Toward this end, we have conducted a systematic mapping. As main result, we contend that formal methods have been barely used to mine crosscutting concerns. Another important contribution is the identification of new research lines. For instance, trying to add formal methods to techniques already stablished for mining crosscutting concerns such as, History-based Aspect Mining [15], Clustering-Based Fan-in Analysis [16] and Concern Mining using Mutual Information over Time [17]. Therefore, there are still different perspectives that could be investigated, aiming at improving the identification of crosscutting concerns.

References

1. Robillard, M., Murphy, G.: Concern graphs: finding and describing concerns using structural program dependencies. In: Software Engineering, 2002. ICSE 2002. Proceedings of the 24rd International Conference on. (may 2002) 406 –416
2. Eaddy, M., Zimmermann, T., Sherwood, K.D., Garg, V., Murphy, G.C., Nagappan, N., Aho, A.V.: Do crosscutting concerns cause defects? IEEE Trans. Softw. Eng. **34** (July 2008) 497–515
3. Petersen, K., Feldt, R., Mujtaba, S., Mattsson, M.: Systematic Mapping Studies in Software engineering. In: EASE '08: Proceedings of the 12th International Conference on Evaluation and Assessment in Software Engineering. (2008)
4. Ubayashi, N., Tamai, T.: Aspect-oriented programming with model checking. In: Proceedings of the 1st international conference on Aspect-oriented software development. AOSD '02, ACM (2002) 148–154
5. Krishnamurthi, S., Fisler, K., Greenberg, M.: Verifying aspect advice modularly. SIGSOFT Softw. Eng. Notes **29** (October 2004) 137–146

6. Nelson, T., Cowan, D., Alencar, P.: Supporting formal verification of crosscutting concerns. In Yonezawa, A., Matsuoka, S., eds.: Metalevel Architectures and Separation of Crosscutting Concerns. Volume 2192 of Lecture Notes in Computer Science., Springer Berlin / Heidelberg (2001) 153–169
7. Thuan, T.N., Ha, N.V.: Using b to verify the weaving of aspects. In: Software Engineering Conference, 2007. APSEC 2007. 14th Asia-Pacific. (dec. 2007) 199 –205
8. Fu, Y., Ding, J., Bording, P.: An approach for modeling and analyzing crosscutting concerns. In: Service Operations, Logistics and Informatics, 2009. SOLI '09. IEEE/INFORMS International Conference on. (july 2009) 91 –97
9. Paula, V.d., Batista, T.: Revisiting a formal framework for modeling aspects in the design phase. In: Proceedings of the Early Aspects at ICSE: Workshops in Aspect-Oriented Requirements Engineering and Architecture Design. EARLYASPECTS '07, Washington, DC, USA, IEEE Computer Society (2007) 6–
10. Dyba, T., Dingsoyr, T., Hanssen, G.K.: Applying systematic reviews to diverse study types: An experience report. In: First International Symposium on Empirical Software Engineering and Measurement. ESEM '07, Washington, DC, USA, IEEE Computer Society (2007) 225–234
11. Pazos-Arias, J.J., García-Duque, J., López-Nores, M.: Locating crosscutting concerns in the formal specification of distributed reactive systems. SIGSOFT Softw. Eng. Notes **30** (May 2005) 1–5
12. Krishnamurthi, S., Fisler, K.: Foundations of incremental aspect model-checking. ACM Trans. Softw. Eng. Methodol. **16** (April 2007)
13. Aaltonen, T., Helin, J., Katara, M., Kellomki, P., Mikkonen, T.: Coordinating aspects and objects. Electronic Notes in Theoretical Computer Science (2003) 248 – 267
14. Kallel, S., Charfi, A., Jmaiel, M.: Using aspects for enforcing formal architectural invariants. Electronic Notes in Theoretical Computer Science (2008) 5–21
15. Breu, S., Zimmermann, T.: Mining aspects from version history. In: Proceedings of the 21st IEEE/ACM International Conference on Automated Software Engineering, Washington, DC, USA (2006) 221–230
16. Zhang, D., Guo, Y., Chen, X.: Automated aspect recommendation through clustering-based fan-in analysis. In: Proceedings of the 2008 23rd IEEE/ACM International Conference on Automated Software Engineering, Washington, DC, USA (2008) 278–287
17. Adams, B., Jiang, Z.M., Hassan, A.E.: Identifying crosscutting concerns using historical code changes. In: Proceedings of the 32nd ACM/IEEE International Conference on Software Engineering - Volume 1, New York, NY, USA (2010) 305–314