

Optimization Approaches for Self-Adaptive Systems

Tim Engbrocks

Institute for Program Structures and Data Organization (IPD)

Advisor: Dipl.-Inform. Martina Rapp

Abstract The complexity of modern software systems is constantly growing. This requires new ways to be found to better manage large scale software systems. Self-Adaptive Systems which autonomously manage themselves using well-defined rules are a solution to this problem. Even though these systems are able to adapt themselves, they struggle with unpredicted changes in their context. This can be solved by using Optimization Approaches for Self-Adaptive Systems which can optimize these systems. There are already many Optimization Approaches, but no way to classify them. This paper proposes a classification for Optimization Approaches for Self-Adaptive Systems.

1 Introduction

The complexity of modern software systems is constantly growing. Most of this growth in complexity stems from the "need to integrate several heterogeneous environments into corporate-wide computing systems, and to extend that beyond company boundaries into the Internet" (Kephart and Chess, 2003 [5]). This has reached a state where the "complexity appears to be approaching the limits of human capability" (Kephart and Chess, 2003 [5]). In combination with the uncertainty about a software systems future operations and environment, that the developers of such complex systems face, it becomes uneconomical to purely operate such a system by human operators.

From this the need for software systems which can autonomously manage themselves arises. In order to achieve this task of autonomous self-management, the system has to be able to:

- detect faults and changes in its environment,
- analyze them,
- decide how to react to them
- and make changes to itself.

To model these abilities Kephart and Chess developed the Monitor, Analyze, Plan, Execute with Knowledge (MAPE-K) feedback loop [5] shown in Figure 1.

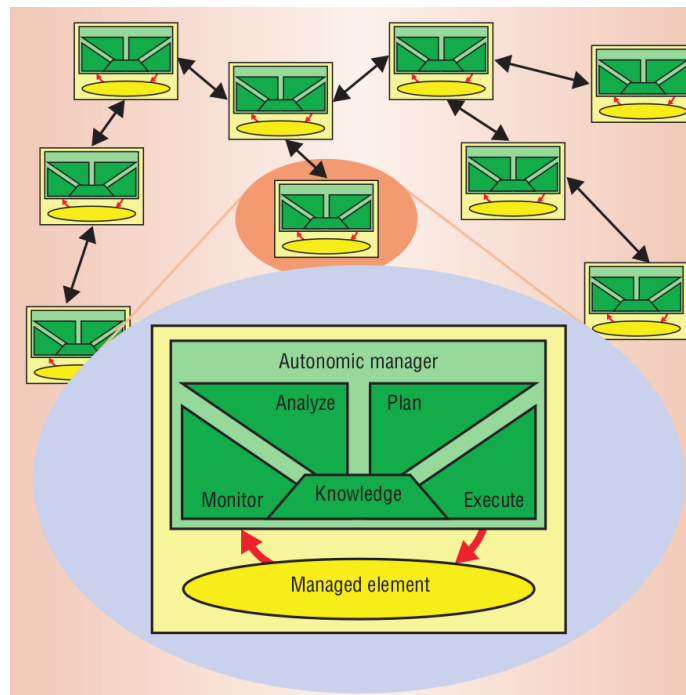


Figure 1: The MAPE-K (Monitor-Analyze-Plan-Execute with Knowledge) feedback loop by Kephart and Chess, 2003 [5]

First the system has to *monitor* itself and its environment. The data, gathered by the monitoring step, has to be *analyzed* to detect changes and faults. If the analyzing step detects, that an adaptation is necessary, the system has to *plan* how to perform the necessary changes. After the changes have been planned, they need to be *executed*. All of this happens with *knowledge* of the environment and the system.

Software systems that can autonomously manage themselves are called Self-Adaptive Systems (SAS) because of their ability to adapt themselves.

To better understand Self-Adaptive Systems, let us take a look at a commonly used example. Imaging you are the system administrator of a large scale online store. This store has four parameters: site traffic, number of purchases, number of active server instances and the number of served advertisements. During your day to day operations you encounter two common tasks:

1. Changing the number of active server instances based on the site traffic.
2. Updating the number of served advertisements based on the number of purchases.

To make your job easier, you decide to use a Self-Adaptive System for these tasks and come up with the following adaptation rules:

1. If the site traffic goes above a certain threshold: Start a new server instance.
2. If the site traffic goes below a certain threshold: Shut down the least used server instance.
3. When the number of purchases decreases: Increase the number of advertisements that are being served on the website.

While Self-Adaptive Systems are better at handling more complex systems, human operators are better at handling uncertainty. This has two reasons. Firstly, the adaptation rules and policies used by SAS are statically created at design time. Secondly, SAS can adapt the software that they are managing but they can not change their adaptation process. Over time this leads to an increasing divergence between the expected results of adaptations and the actual results, when the environment changes in ways that were not predicted by developers during the design time of the system.

This can be illustrated by the previous example. Imagine that the Self-Adaptive System has been in operation for some time and you collected data on the systems performance. You notice that the number of start/shut down operations per server instance has increased which has led to an increased electricity bill. After further analyzing the data you discover that the site traffic often fluctuates around the threshold that you set. Each time that the site traffic crossed this threshold a new server instance was started or the least used instance was stopped. As a human operator you would have waited on how the site traffic develops before starting or shutting down server instances. The SAS could not handle this type of uncertainty and only reacted to the current state of its environment.

Optimizations are necessary to increase the performance and effectiveness of Self-Adaptive Systems. There are already many approaches on how to optimize SAS. Some of them focus on updating adaptation rules and policies during the systems runtime. Others dynamically change at which level of the system adaptations are performed.

The previous example could benefit from such optimizations by dynamically updating adaptation rules during times of the day which have historically seen large fluctuations around the thresholds that were set.

While there are already many Optimization Approaches (OA) for SAS, there is no classification for them. Because of this, the existing OA can not be easily compared and it is difficult to identify areas which require further research. This paper aims to provide such a classification for Optimization Approaches for Self-Adaptive Systems.

To derive a classification for Optimization Approaches for Self-Adaptive Systems, chapter 2 will first explain how SAS are classified using three different approaches. Based on these approaches, a classification for OA for SAS will be derived and proposed in chapter 3. In chapter 4 the proposed classification will be applied to some existing OA. Lastly, chapter 5 will finish with a conclusion and recommendations for future research directions.

2 Classification of Self-Adaptive Systems

There are different approaches on how to classify and describe Self-Adaptive Systems which all focus on different usages. The three approaches that will be highlighted by this paper are:

- FORMS from Weyns et al., 2012 [3]
- Berns and Ghosh, 2009 definition of self-* properties [2]
- Krupitzer's et al., 2015 taxonomy for Self-Adaptive Systems [6]

These approaches contain ideas that will be used to construct the classification for Optimization Approaches for Self-Adaptive Systems in the next chapter.

In their 2012 paper FORMS [3] Weyns et al. propose a formal reference model for describing Self-Adaptive Systems. The goal of FORMS is to provide a well defined basis for talking and reasoning about Self-Adaptive Systems. This is achieved by providing a definition of FORMS which classifies Self-Adaptive Systems using Z notation. By composing a Self-Adaptive System from different levels of components called subsystems, where each level can adapt the level beneath, FORMS models the self-adaptive process using reflection. The bottom level is populated by base components and domain models. On top of these base components are reflective components and reflection models. The relation between these component or primitives is depicted in Figure 2.

The concept of different layers of reflection will be useful later to distinguish Optimization Approaches for Self-Adaptive Systems from Self-Adaptive Systems.

Although some of the first papers on Self-Adaptive Systems, for example, "The Vision of Autonomic Computing" by Kephart and Chess [5] focused on self-adaptation, there are other aspects of software systems that can benefit from the ideas introduced by Self-Adaptive Systems. Berns and Ghosh, 2009 [2] identified and described such aspects which they call self-* properties. These self-properties can be seen in Figure 3. They are useful to:

- Communicate the abilities and goals of a system.
- Establish well-defined goals for the system.

These self-* properties show that there are aspects besides the management of software which can benefit from the ideas of Self-Adaptation. This will be useful for the proposed classification as it gives a clear setting for the Optimization Approaches to work within. The taxonomy for Self-Adaptive Systems by Krupitzer's et al., 2015 [6] in Figure 4 is based upon the 5W+1H questions by Salehie and Tahvildari, 2009 [7]. These questions are: Where, When, What, Why, Who and How. Each of these questions is responsible for a different aspect of Self-Adaptive Systems and corresponds to a dimension of the taxonomy.

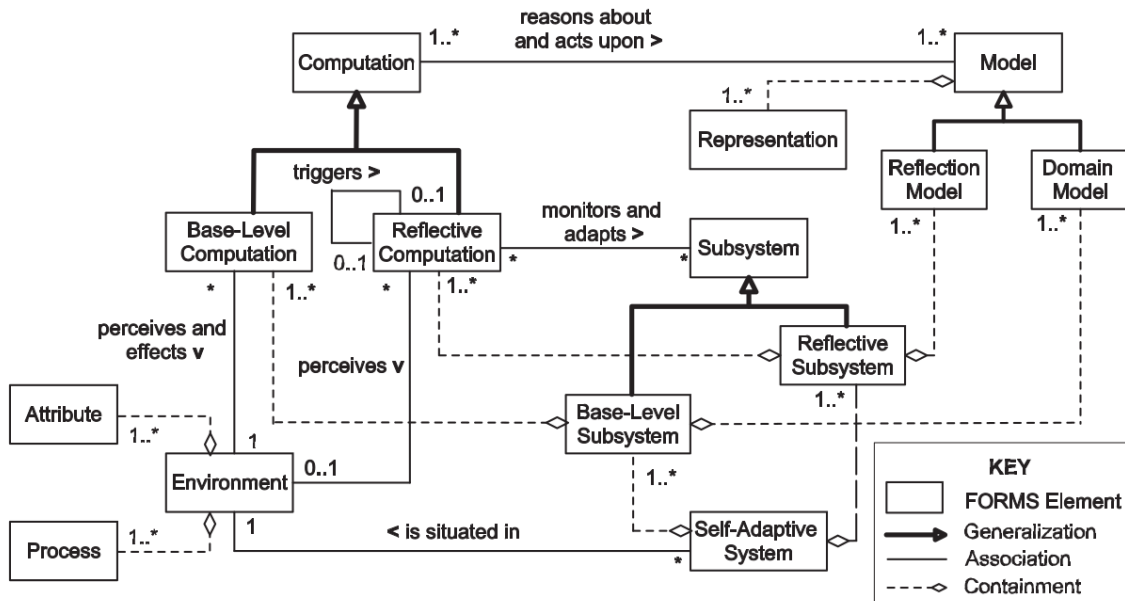


Figure 2: FORMS primitives by Weyns et al., 2012 [3]

SUMMARY OF SELF-* PROPERTIES. F = SET OF EXTERNAL ACTIONS, P = PREDICATE OVER GLOBAL STATE.

Self-* Property Name	Adversarial actions	Behavior	Predicate maintained
Self-stabilization	All transient failures	Restore	P
Self-adapting	Change of environment R	Restore	if R_i then P_i
Self-healing	Some $C \subseteq F$	Restore	P
Self-organizing	Process join or leave	Maintain, improve, or restore	P
Self-protecting	Some malicious actions	Maintain	Predicate over trust
Self-optimization	Some $C \subseteq F$	Improve or maximize/minimize as appropriate	An objective function of the global or local state
Self-configuration	Some $C \subset F$, often includes user demands for service	Maintain, improve, or restore using configuration changes	Predicate over system configuration to optimize performance. Sometimes addresses geometric invariant
Self-scaling	Changes of system scale or demand	Restore, or improve	Predicate over service quality
Self-immunity	Some $C \subseteq F$	Eventually maintain	P
Self-containment	Some malicious actions	Maintain (for a subset of processes)	Predicate over trust

Figure 3: self-* properties by Berns and Ghosh, 2009 [2]

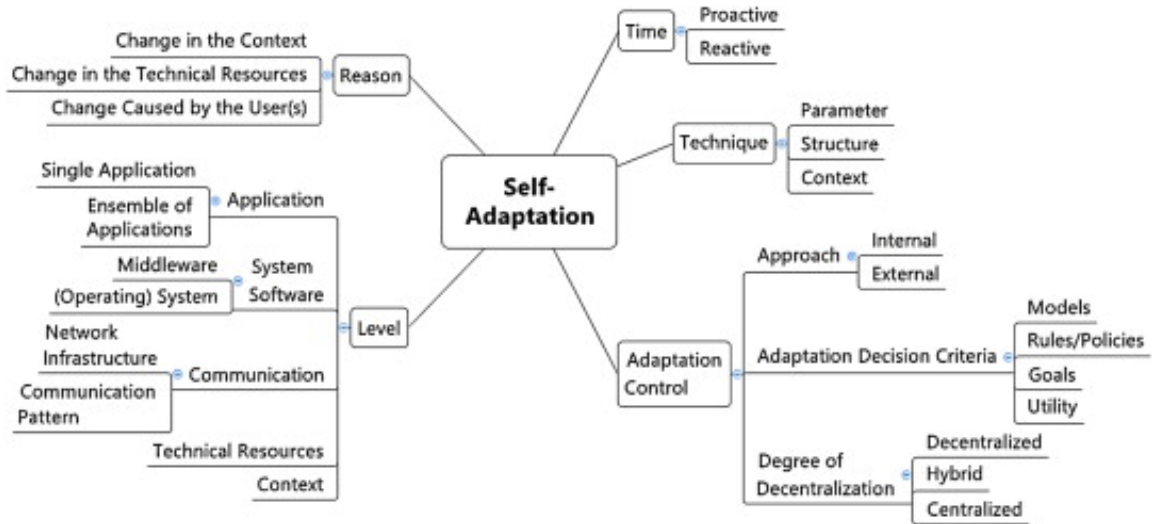


Figure 4: Taxonomy for Self-Adaptive Systems by Krupitzer et al., 2015 [6]

Why First there needs to be a reason for a Self-Adaptive System to adapt. Why an adaptation should be performed is answered by the Reason dimension. According to the taxonomy reasons for an adaptation can be changes in either the context, a technical resource or changes caused by the user.

Where The question of where asks on which level of the system changes need to occur. The different levels on which changes can occur include:

- Different levels of applications from the operating system to a user application.
- How systems communicate with each other.
- The technical resources that are needed by the system.
- The context in which the system operates.

When While the original When-Question by Salehie and Tahvildari tries to understand all temporal aspects of Self-Adaptive Systems, including how frequently changes should occur and if they happen continuously, the taxonomy only answers the question when changes should be performed. For this purpose the Time dimension differentiates between systems that perform changes proactively or reactively.

What In addition to the question of where and when changes should occur, it is also important to know what changes should occur. There are different techniques that can be used. The Technique dimension of the taxonomy differentiates between systems that change parameters, their structure or their context.

Who After answering where, when, what and why changes should be performed, it is necessary to select who is responsible for these changes. According to Salehie and Tahvildari it is also important to establish if the changes can be performed fully autonomous or if the involvement of human operators is necessary. The taxonomy does not directly address all of these concerns but states that: "N/A (nature of a SAS leads to an automatic type of adaptation)" (Krupitzer et al., 2015 [6]).

How Lastly, after determining the where, when, what, why and by who, there needs to be a way to perform the required changes. This is answered by asking how the changes should be performed and corresponds to the Adaptation Control. The three main factors of the Adaptation Control are the degree of decentralization, the adaptation decision criteria and the approach taken by the system, which divides Self-Adaptive Systems into those where the adaptation logic is part of the application logic and those with separated adaptation and application logic.

After classifying Self-Adaptive Systems the following question can be asked: Which parts of a Self-Adaptive System can be optimized? To answer this we will start by looking at which parts can not be optimized or do not benefit from optimization.

The first part that can not be optimized is the environment which provides the Reason dimension. While the environment for a Self-Adaptive System can be chosen in a way which is most beneficial for the system and can be influenced by actors, the behavior of the systems environment can generally not be controlled completely.

Another dimension of Self-Adaptive Systems that can not be optimized, or is not useful to optimize, is the Time dimension. This dimension is mostly a design decision on how the system should behave and be constructed. It is also a question of how to handle uncertainty and the level of accepted risk. A proactive system can prevent faults and degradation in Quality-of-Service metrics, but it can also predict the wrong changes which

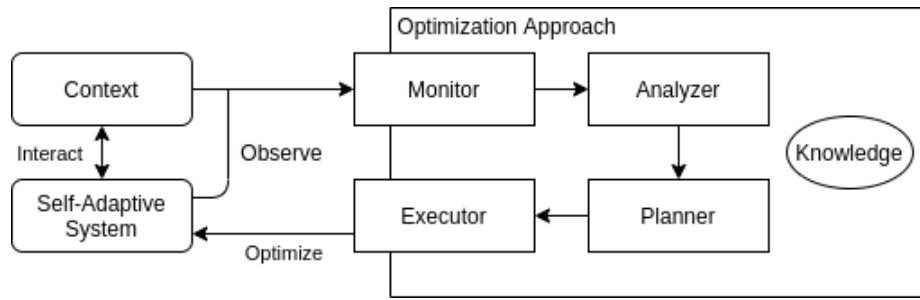


Figure 5: Mapping the optimization process to the MAPE-K feedback loop.

can lead to a situation where the system itself generates faults by reacting in a way that is contradictive to its goals. A reactive system can not prevent faults like a proactive system, but its behavior can be much more stable, because it only has to react to a change and not predict that change as well.

Lastly, the question of who is responsible can not be optimized, because the taxonomy simply answers it, by referring to the name "Self-Adaptive System" which implies that the system itself is responsible for managing adaptations.

The three remaining dimensions can be optimized or can benefit from being adapted dynamically. These are the Adaptation Control, the Level and the Technique.

The Approach and the Degree of Decentralization used by the Adaptation Control can not be optimized because they are design decisions of how the system is built. However, the Adaptation Decision Criteria can be optimized. An optimization of the Adaptation Decision Criteria could, for example, be to dynamically adapt the rules and policies at runtime to better reflect a changing environment.

Another dimension that can be optimized is the Technique. This can be optimized by changing what gets adapted by the system.

The last dimension that can be optimized is the Level, which can be done by dynamically changing at which level of the system adaptations should be performed.

This explanation of how Self-Adaptive Systems are classified, leads us to how we can classify Optimization Approaches for Self-Adaptive Systems. The main concepts that will be used for the classification of Optimization Approaches are:

- The concept of reflectiveness as a way to model self-adaptation.
- The 5W+1H questions to determine how a Self-Adaptive System operates.
- The MAPE-K feedback loop a general base for understanding Self-Adaptive Systems.
- The concept of self-* properties which enables us to differentiate between self-adaptation and other self-* abilities.
- A taxonomy for Self-Adaptive Systems which will be the basis for the proposed classification of Optimization Approaches, because Self-Adaptive Systems are so closely related to their Optimization Approaches.

3 Proposal for Classification of Optimization Approaches

Using the idea from Weyns et al. 2012 paper FORMS [3] to compose Self-Adaptive Systems from layers of base and reflective components, an Optimization Approach for Self-Adaptive Systems can be thought of as a reflective component in a layer above the Self-Adaptive System.

This allows the application of multiple Optimization Approaches and even the optimization of Optimization Approaches.

It also establishes a connection between Optimization Approaches and Self-Adaptive Systems. This leads to the realization that a classification for Optimization Approaches of Self-Adaptive Systems should be based upon the same principles as the classifications for Self-Adaptive Systems.

Another benefit is a clear distinction between Self-Adaptive Systems and Optimization Approaches for Self-Adaptive Systems. Self-Adaptive Systems and Optimization Approaches are both composed of reflective components and not base components.

Yet the Self-Adaptive System adapts base components, while the Optimization Approach adapts other reflective components.

Just like the adaptation process of Self-Adaptive Systems can be understood using the MAPE-K (Monitor-Analyze-Plan-Execute with Knowledge) feedback loop by Kephart and Chess, 2003 [5]. The process of optimizing a Self-Adaptive System can also be expressed using the MAPE-K feedback loop:

- Firstly the optimization approach has to constantly monitor the Self-Adaptive System and its context.
- The data gathered from monitoring can then be analyzed to decide whether or not an optimization is necessary and what should be optimized.
- After deciding that something should be optimized, there needs to be a plan on how to optimize.
- Lastly the planned optimization can be executed.
- During all of these steps the optimization approach requires knowledge about the Self-Adaptive System and its context.

Following Krupitzer's et al, 2015 [6] taxonomy for Self-Adaptive Systems, a classification for Optimization Approaches of Self-Adaptive Systems should answer the 5W+1H questions by Salehie and Tahvildari, 2009 [7]. These questions are: Where, When, What, Why, Who and How?

Figure 6 shows the proposed classification for Optimization Approaches for Self-Adaptive Systems. The classification consists of six dimensions: Location, Time, Purpose, Approach, Technique and Reflective Adaptation Control. Each dimension is related to one of the 5W+1H questions.

Location Where in the system is an optimization necessary?

There are three main parts in a Self-Adaptive System that can require optimizations during its lifetime. These are:

- The Adaptation Control which can, for example, be optimized by changing or generating new adaptation rules.
- The Level on which adaptations are performed.

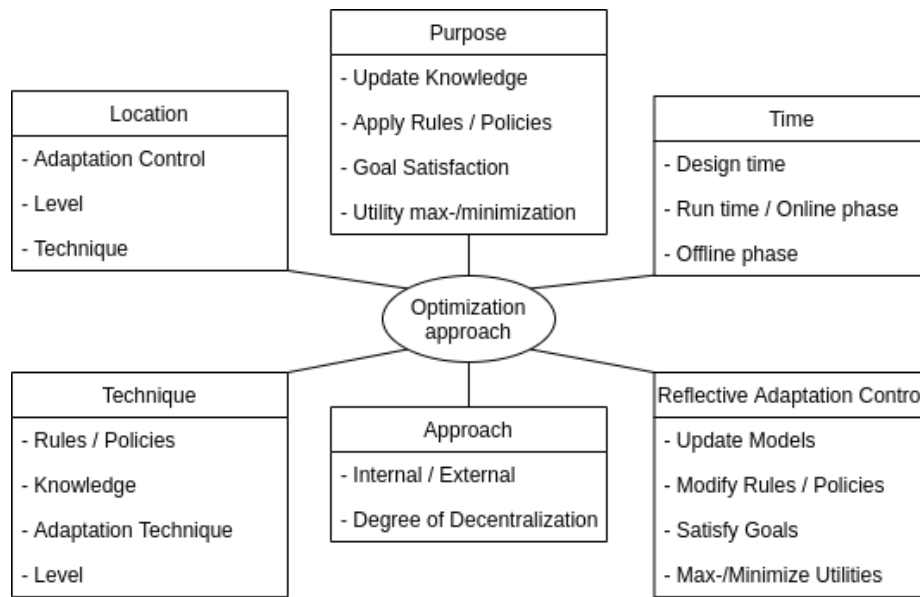


Figure 6: The proposed classification for Optimization Approaches for Self-Adaptive Systems

- The Technique that is used to perform adaptations.

Time When are optimizations performed?

Similarly to Self-Adaptive Systems, Optimization Approaches can be differentiated by comparing when they perform optimizations. There are three different phases during the lifetime of a Self-Adaptive System where optimizations can occur. These are:

- At the runtime of the system, which can also be called the online phase.
- During the design time of the system.
- While training the system or during its offline phase.

The training phase can happen in parallel to the run and design time of the system. Examples for this would be when the initial parameters for a Self-Adaptive System are chosen by training a domain model or when generating a new model for the system by training an updated domain model parallel to the running system.

Technique What gets changed to perform the optimization?

What gets changed in a Self-Adaptive System relates to where an optimization is necessary. Thus, the following parts can be changed:

- The Rules / Policies used by the Adaptation Decision Criteria.
- The Knowledge of the system, which for example can consist of domain models.
- The Adaptation Technique used by the Self-Adaptive System.
- The Level on which the Self-Adaptive System performs changes.

One could argue that the Goals or Utility functions of a Self-Adaptive System could also be changed to optimize its performance. In most cases this is not advisable because Goals and Utility functions can encode information like legal or safety parameters, of which the system has no intrinsic knowledge.

Purpose Why should an optimization occur?

This is perhaps the most important question for any Optimization Approach as it determines if an optimization is necessary. There can be several reasons for performing an optimization in a Self-Adaptive System:

- The system learned new information about, for example, its environment and needs to update its domain models.
- Due to of external or internal changes, a Goal might not be satisfiable anymore.
- The difference of the actual versus the expected outcome of a rule / policy has exceeded a threshold.
- A Utility function needs to be maximized or minimized.

Approach Who is responsible for performing the optimization?

The original "Who?" question by Salehie and Tahvildari focuses on the level of human intervention that is need by the system. Self-Adaptive Systems have no need for human intervention, because they operate on clearly defined rule sets. It is in the nature of Optimization Approaches for Self-Adaptive Systems to change these rule sets. This can lead to problems that make human intervention necessary. A result of this is that, in contrast to the the taxonomy for Self-Adaptive Systems, the Approach for Optimization Approaches should be its own dimension.

- Is there an internal component that performs changes or are they performed by an external actor.
- Which degree of decentralization is used? Is each component responsible for itself (fully decentralized), are all components optimized by a central entity (fully centralized) or is a hybrid approach used?

Reflective Adaptation Control How is the optimization applied to the system?

Just like the Self-Adaptive System can be described by its Adaptation Control, which is responsible for applying changes to the system, an Optimization Approach can also be described by how it applies its optimizations to the system. In reference to FORMS by Weyns et al., 2012[3], which uses reflective operations to change components in the system, the Adaptation Control of the optimization approach will be called Reflective Adaptation Control. The Reflective Adaptation Control can be classified by the following behaviors:

- Models: The system updates its models after receiving new information about its models.
- Rules / Policies: If the system detects, that it is in a state for which an adaptation rule/policy exists, it should perform that adaptation.
- Goals: Adaptation should be performed if the system does not meet pre-defined goals.
- Utility: An adaptation occurs to maximize or minimize a utility function.

Because the proposed classification is based upon the MAPE-K feedback loop and the 5W+1H questions, it is useful to understand how they relate to each other. This is depicted in Figures 7 and 8.

Optimization Approach	MAPE-K
Location	Monitor, Analyze, Execute
Time	Monitor, Execute
Technique	Plan, Execute
Purpose	Analyze
Approach	Plan, Execute
Reflective Adaptation Control	Plan, Execute

Figure 7: How the MAPE-K feedback loop by Kephart and Chess, 2003[5] relates to the dimensions of the classification for Optimization Approaches for Self-Adaptive Systems.

Optimization Approach	5W+1H questions
Location	Where
Time	When
Technique	What
Purpose	Why
Approach	Who
Reflective Adaptation Control	How

Figure 8: How the 5W+1H questions by Salehie and Tahvildari, 2009 [7] relate to the dimensions of the classification for Optimization Approaches for Self-Adaptive Systems.

4 Classifying Existing Optimization Approaches

FUSION by Elkhodary et al, 2010 [1] uses a learning and an adaptation cycle. The adaptation cycle corresponds to the Self-Adaptive part of the system and does not directly affect the learning cycle. The learning cycle is the optimization approach used by FUSION, it changes the behavior of the adaptation cycle. Both of these cycles continuously run in parallel to each other.

- Location: FUSION optimizes the Adaptation Control.
- Time: FUSION uses both the design and run time to perform optimizations. The design time is used to generate an initial model for the learning cycle. During the run time the learning cycle monitors adaptations and updates the Knowledge Base to improve future adaptations.
- Technique: FUSION updates the rules used by the Self-Adaptive System and the shared knowledge base.
- Purpose: FUSION aims to decrease the difference between the expected and actual outcome of adaptation rules.
- Approach: FUSION acts as an external actor.
- Reflective Adaptation Control: FUSION applies optimizations by updating models and modifying rules.

Learning Revised Models for Planning in Adaptive Systems by Sykes et al., 2013 [8] uses "a non-monotonic probabilistic rule learning technique, NoMPRoL, that finds hypotheses consisting of new rules specifying the probability of an action achieving its specified outcome under particular conditions" (Sykes et al., 2013 [8])

- Location: The optimizations focus on the Adaptation Control.
- Time: The run time is used to perform optimizations.
- Technique: This approach optimizes the adaptation rules and domain models.
- Purpose: The optimizations aim to generate new and better adaptation rules and domain models.
- Approach: The optimizations are performed by an internal component that sits between the domain model and the Self-Adaptive System.
- Reflective Adaptation Control: This approach generates new adaptation rules.

FloT by Moraes do Nascimento and Pereira, 2016 [4] is a framework for using agent-based systems to construct self-organizing systems for the Internet of Things.

- Location: FloT optimizes three different levels: The physical, communication and application layer.
- Time: FloT performs its optimizations during the run time of the system.
- Technique: FloT changes the Level, Adaptation Technique and Rules / Policies of a Self-Adaptive System.
- Purpose: Generate self-organizing behavior in a Self-Adaptive System.
- Approach: FloT uses a decentralized agent-based approach.
- Reflective Adaptation Control:

5 Conclusion

This paper proposed a classification for Optimization Approaches for Self-Adaptive Systems. The classification was derived from the same principles as the taxonomy for Self-Adaptive Systems by Krupitzer et al., 2015 [6]. The main principles that were used to derive the classification are:

- The 5W+1H questions by Salehie and Tahvildari, 2009 [7].
- The MAPE-K feedback loop which was created by Kephart and Chess, 2003 [5].

To classify Optimization Approaches the proposed classification provides six dimensions: Location, Time, Purpose, Approach, Technique and Reflective Adaptation Control. These are shown in Figure 6.

The Location dimension asks on which level optimizations in a Self-Adaptive System might be necessary. Time distinguishes between Optimization Approaches that are performed during the design time of the system, during the run time or online phase of the system or approaches that optimize systems during their offline phase. The dimension of Purpose provides a reason for Optimizations to be performed. Approach determines who is responsible for optimizations and Technique states which parts of a Self-Adaptive System can be optimized. Lastly, the Reflective Adaptation Control classifies how optimizations are performed.

After proposing a classification, it was exemplarily applied to several existing Optimization Approaches. This is a part where further research is required. The classification has to be applied to further existing approaches to verify its validity and completeness. This would also yield a way to determine areas which require further research.

References

- [1] Naeem Esfahani Ahmed Elkhodary and Sam Malek. “FUSION: A Framework for Engineering Self-Tuning Self-Adaptive Software Systems”. In: *Proceedings of the Eighteenth ACM SIGSOFT International Symposium on Foundations of Software Engineering*. FSE '10. Santa Fe, New Mexico, USA: Association for Computing Machinery, 2010, pp. 7–16. ISBN: 9781605587912. DOI: 10.1145/1882291.1882296.
- [2] Andrew Berns and Sukumar Ghosh. “Dissecting Self-* Properties”. In: *2009 Third IEEE International Conference on Self-Adaptive and Self-Organizing Systems*. 2009, pp. 10–19. DOI: 10.1109/SASO.2009.25.
- [3] Sam Malek Danny Weyns and Jesper Andersson. “FORMS: Unifying Reference Model for Formal Specification of Distributed Self-Adaptive Systems”. In: *ACM Trans. Auton. Adapt. Syst.* 7.1 (May 2012). ISSN: 1556-4665. DOI: 10.1145/2168260.2168268.
- [4] Nathalia Moraes do Nascimento and Carlos José Pereira de Lucena. “FloT: An agent-based framework for self-adaptive and self-organizing applications based on the Internet of Things”. In: *Information Sciences* 378 (2017), pp. 161–176. ISSN: 0020-0255. DOI: 10.1016/j.ins.2016.10.031.
- [5] J.O. Kephart and D.M. Chess. “The vision of autonomic computing”. In: *Computer* 36.1 (2003), pp. 41–50. DOI: 10.1109/MC.2003.1160055.
- [6] Christian Krupitzer et al. “A survey on engineering approaches for self-adaptive systems”. In: *Pervasive and Mobile Computing* 17 (2015). 10 years of Pervasive Computing’ In Honor of Chatschik Bisdikian, pp. 184–206. ISSN: 1574-1192. DOI: 10.1016/j.pmcj.2014.09.009.
- [7] Mazeiar Salehie and Ladan Tahvildari. “Self-Adaptive Software: Landscape and Research Challenges”. In: *ACM Trans. Auton. Adapt. Syst.* 4.2 (May 2009). ISSN: 1556-4665. DOI: 10.1145/1516533.1516538.
- [8] Daniel Sykes et al. “Learning revised models for planning in adaptive systems”. In: *2013 35th International Conference on Software Engineering (ICSE)*. 2013, pp. 63–71. DOI: 10.1109/ICSE.2013.6606552.