A Proactive Self-Adaptive Framework for Context-Aware Mobile Applications

Md. Shafiuzzaman
Institute of Information Technology, University of Dhaka

Amit Seal Ami
Institute of Information Technology, University of Dhaka

Rayhanur Rahman
Institute of Information Technology, University of Dhaka

ABSTRACT

Today’s mobile applications are continuously renovating user experiences by processing underlying contextual information of the hand-held devices. The mobility provided by these devices is encouraging users to install and use applications on the move. As a result, underlying context of these devices keep changing constantly. As the changes in context are beyond control of the application, a context-aware application needs to be self-adaptive in the sense that it can update its services dynamically in runtime to meet overall user requirements. With the remarkable technological advances of mobile devices, collecting contextual information is no longer an issue. However, adapting the contextual changes are still an open problem as most of the existing self-adaptive frameworks are not proactive in nature and the existing proactive frameworks suffer for their computational complexities as they rely on architectural reconfiguration. As a result, none of those have become effective enough for mobile applications. In order to tackle the aforementioned challenges, this paper proposes a light-weight proactive solution that suggests to offer one or more alternative features of each functionality so that the affected feature can be disabled or reconfigured with another mutually exclusive feature depending on the contextual requirements. Several experiments were conducted to test the achieved context-awareness as well as the imposed overhead. The tests were conducted over battery consumption as it is the most common context for most of the mobile applications. The proposed framework extended the battery life 30.46% than the non-adaptive application. The activity of stopping and activating features had no significant impact on the battery life as only 1.5% battery was consumed when we deactivated 7 features at once. The time required to stop and active features increased linearly with the number of features though it was in the range of milliseconds. It took 0.6 seconds to resume the application after deactivating 7 features simultaneously. The proposed approach brings about an innovation for solving the aforementioned challenges. It extracts analytical model from software requirements models to estimate the probable execution flow and assess the context requirements of running application against the ideal requirements mentioned in requirements model. According to the assessment result, it updates the feature set available for the users. The novelty of this approach is to use of requirements model in feature-oriented adaptation to provide adaptivity. where and anytime. At the same time, such systems require strict reliability requirements in the sense they never go out of service. But the traditional mobile applications fail to provide uninterrupted service during frequent contextual changes as they are not non-adaptive in nature. A self-adaptive system is characterized with the capability of re-configuring a software system at runtime to meet expected Quality of Service (QoS) [9]. In practice, engineers try to ensure this capability by explicitly programming alternative behaviors and implementing heavy exception handling mechanism [7]. This is quite unrealistic to ensure exception handling for each combination of different contexts. In addition, using this approach, the architecture and the code become unmanageable as the alternative behaviors cannot be kept separated from each other [7].

In literature, quite a few approaches [6,8,9,11] can be found that target the challenges of self-adaptive systems. Initial approaches were mostly confined into ensuring adaptivity by updating configuration parameters dynamically at runtime [11,12]. These parameterized adaptation became popular for its simplicity but failed to support large-scale adaptation [9]. Other approaches can be divided into two broader domains - (i) Architecture-based adaptation [6,8] and (ii) Feature-oriented adaptation [9,17]. Architecture-based adaptation depends on adding or replacing architectural components at runtime. On the other hand, feature-oriented adaptation does not require knowledge about internal structure of the software system instead it pulls out the affected feature when it goes out of service and restore the system with an alternative feature simultaneously. Both of the approaches failed to fulfill required stuffs of self-adaptation for mobile applications. Searching for an optimal configuration at the architectural-level is computationally very expensive for mobile applications. In fact, many architecture-based optimization algorithms are shown to be NP-hard [21,22]. On the other hand, feature-oriented adaptation suffers from proactive decision making as it lacks of knowledge about the execution flow. A mobile application becomes unusable if context changes cannot be adapted before reaching failure state [8].

Keywords
Context awareness, self-adaptive system, mobile application, feature-orientation, model-driven development

1. INTRODUCTION

Mobile software systems are characterized by a high degree of unpredictability and dynamism as those can be installed and used any-
Extraction of analytical model from traditional software requirements models is formalized in our another contribution [1] where we propose a method of modeling and representing context-aware self-adaptive software systems by extending the ‘Model-Driven Development’ framework [2]. This paper extends that approach to model the overall system as a workflow of some abstract functionalities while each of those can be implemented in one or more alternative ways. Then, all these alternatives merged into a single Markov Decision Process (MDP) [15] with a finite state automaton where each state of the automaton represents an alternative feature of an abstract functionality and the paths represent possible execution flows denoting the probabilities of execution of those states. At runtime, the system is treated as a variable so that features can be enabled and disabled dynamically on availability of contextual factors. To automate this process, an interpreter navigates the automaton developed at design phase state by state and offers the best possible set of features to the users depending on the availability of the contextual requirements.

The proposed framework is evaluated by identifying the overall usability increased by the framework. Besides, we also test the runtime overhead imposed by the framework. This is done by varying number of abstract functionalities and number of alternative features suggested by [21]. The remainder of this paper is organised as follows: Section 2 describes the existing works of this field. Section 3 presents the framework in detail. Section 4 validates the performance of the framework. Finally, section 5 concludes the paper with future work directions.

2. RELATED WORK

Over the past decade, researchers have contributed with different methodologies, techniques, strategies and frameworks to engineer the adaptation logic. In this section, we provide most notable and relevant approaches to our work.

Bennani et. al [11] propose a parameterized model that reconfigures system’s behavior at runtime by measuring system characteristics (e.g., workload). Ryutov et al. [12] also use parameterization to develop adaptive access control and trust negotiation framework. Though parameterized models were welcomed for their simplicity, they lack to support large-scale adaptation [9]. There are several noteworthy approaches that stand on structural changes, such as adding, removing and replacing software components [6], changing system’s architectural style [13] and rebinding component’s interfaces [14]. This paradigm of adaptation strategy is commonly referred as architecture-based adaptation. The most recent and notable approach of this domain is proposed in [8]. It proposes a three-layer (Goal management, change management and component control) reference model to manage structural changes. Usually, these approaches last from design time to runtime. At design-time, architectural representation of the system is commonly referred as architecture-based adaptation. The most recent and notable approach of this domain is proposed in [8].

Enabling and disabling features at runtime to reconfigure the system is also a popular paradigm of self-adaptive software development research. This paradigm is commonly called as ‘Feature-Oriented Adaptation’ [17]. A recent solution of this paradigm [9] proposes a learning-based framework for run-time management of feature selection and deselection. This paradigm is computationally less expensive than architecture-based approach and also popular for simplicity, though it has less accuracy on taking adaptation decision as it is not aware of any architectural parameters [7] as well as it can not determine a potential failure proactively [18].

3. APPROACH OVERVIEW

A software system is built on several user requirements and an user requirement can be interpreted as several sequential user activities. Each of those activities can have several different implementations. Traditionally, while designing an application, software engineers choose an implementation approach by making a trade-off between cost and user expectations. Therefore, traditional software systems provide a static set of features. In contrast, this framework proposes to implement a functionality in several alternative ways so that a feature can be turned on or off in running application depending on the condition of contextual requirements. As a result, users may not deal with a similar pool of features each time while launching the application but they will never face any service interruption until complete shutdown of the device. Users must be aware of this dynamic behavior of the application through Service Level Agreement (SLA). To implement this dynamism, we propose to make use of Model-Driven Development framework [5]. In this development framework, software engineers model a feature targeting the requirements and map it to a part of the software system. We make use of those requirements models in our framework by injecting the contextual factors into them and also keeping them alive at runtime to assess any QoS violation. For this purpose, the requirements models are transformed to analytical models using the transformation approach described in [23] which are used to verify the running application against the requirements models at runtime. Using this verification result, the adaptation then adapts itself by re-configuring the available feature sets. Injecting self-adaptivity with applications requires efforts in both design and implementation phase of the development life-cycle. The design phase modeling is based on generating feature selection and modeling strategy while runtime adaptation relies on context collection, analysis and execution of dynamic feature selection. Following sub-sections describe these two phases of the framework in detail.

3.1 Design time

To employ self-adaptability in a running application software engineers are required to perform some additional task while designing the application. This starts with modeling the user requirements with some alternative features. It also modifies the conventional requirements models to make space for integrating contextual factors within the model. Finally, it outputs an analytical model which will be used by the running system to make decision about adaptation. The design phase of the framework consists of following steps:

(1) Feature Modeling
(2) Feature Priority-Adaptability Mapping
3. Context Injection
(4) Analytical Modeling

3.1.1 Feature Modeling. The input of this feature modeling process is a set of abstract functionalities

\[ A = \{A_1, A_2, \ldots, A_n\} \]

which are identified as a part of requirements engineering process. These abstract functionalities are processed through following workflows to model more than one concrete implementations for each of them.

3.1.1.1 Identification of the Contextual Factors. At design time, contexts can be captured by analyzing each of the abstract functionalities separately. For this reason, engineers need to use their domain knowledge or their past experiences on observed behavior of similar functionalities. The set of contextual factors can be formally notated as \( C \). We have partitioned this set into two disjoint sets:

— Static context, \( C_S \)
— Dynamic context, \( C_D \)

\( C_S \) captures the set of stable contexts, which will not change later at runtime. \( C_S \) may include the known physical constraints that regulate certain environmental phenomena, \( C_S \) must be considered while designing the system as some architectural decision may depend on it. \( C_D \) instead captures the contextual factors that are likely to change over time. \( C_D \) must be monitored at runtime to check if they are available with their minimum requirements. We exclude \( C_S \) from our framework as they are not likely to change at runtime where \( C_D \) should be considered for their dynamic characteristics at runtime. Therefore, we inject \( C_D \) into the requirements models. As only the \( C_D \) are considered in our framework, we will describe \( C_D \) as \( C \) throughout this paper. So, we may define \( C_D \) as

\[ C = \{C_1, C_2, \ldots, C_n\} \]

Here, each \( C_i \in C \), \( 1 \leq i \leq n \) signifies an unique context which characterizes the runtime behavior of each abstract functionality. The relation between the contexts and the abstract functionalities can be formally defined as equation \( 1 \)

\[ A_i \in A, \mu : P(C) \rightarrow A \quad (1) \]

where \( 1 \leq i \leq n \) and \( P(C) \) denotes the power set of contextual factors and \( \mu \) represents the combine effect of \( P(C) \) on each abstract functionality. Here, the power set of contextual factors is used as every context may not affect each of the functionalities.

3.1.1.2 Determination of the Levels of Contextual Factors. Level is a scale of amount that is usually used to define the categories of any measurement. Engineers need to determine the levels of each of the contextual factors as the alternative features will be modeled targeting each of those levels later on. Here, we propose to use relative measurement instead of actual values as the real context values are difficult to measure and also device dependent. On the other hand, the relative values can be determined easily by comparing the minimum requirements of each of the alternative features. Furthermore, these relative levels can be updated easily later on by obtaining real data from the running system. For this reason, the proposed framework defines the levels of the contextual factors as the subset of Natural Numbers, \( N \). This is formally defined as \( 2 \)

\[ l_i \in N, \quad 1 \leq i \leq n \quad (2) \]

So, a level of context lies between \( 0 \leq l_i \leq n \). Level of a context becomes ‘0’ when a context is insensitive to a feature.

Listing 1: Implementations of two alternative features

\[ \text{@Feature (name = "F1")}
\]
\[ \text{@Level (factors = \{"C1", "C2", "C3"\}, values = \{0,1,2\})}
\]
\[ \text{implementation_F1 (input) {}
\]
\[ \text{ // Implementation}
\]
\[ \text{}}
\]
\[ \text{@Feature (name = "F2")}
\]
\[ \text{@Level (factors = \{"C1", "C2", "C3"\}, values = \{1,2,3\})}
\]
\[ \text{implementation_F2 (input) {}
\]
\[ \text{ // Implementation}
\]
\[ \text{}}
\]

3.1.2 Feature Priority-Adaptability Mapping. In this step, engineers need to design one or more alternative features for each of the abstract functionalities targeting the different levels of the contextual factors. Therefore, engineers need to model a feature for each \( \mu \) defined in equation \( 1 \). The relation between contextual factors, abstract functionalities and their alternative features can be formalized using equation \( 3 \). Here, \( F \) denotes the set of features of an abstract functionality,

\[ F = \{F_1, \ldots, F_n\} \]

such that each of the members of this set implements the abstract functionality targeting each \( \mu \).

\[ F, C \models A_i \in A \quad (3) \]

That is, a \( F \) ensures satisfaction of an abstract functionality \( A_i \in A \) in the context of \( C \).

To visualize the abstract view of a feature, we prepare a template for expressing the alternative features. For this purpose, we use the ad-hoc annotation @Functionality to express the functionality, @Feature to define each of the features of that functionality and @Level to express the levels of the contexts. An example template of expressing the feature modeling is illustrated in Listing \( 1 \).

3.2 Feature Priority-Adaptability Mapping. Frequent updates of available feature sets at runtime may create extra overhead on context information acquisition and processing for decision making. To tackle this issue, the framework presents a trade-off that updates the feature sets considering their priority and importance. For this reason, it proposes to make use of Feature Prioritization Policy \( 25 \). Prioritization relies on giving more weight to most important features. Following requirements engineering approach \( 25 \), we have defined a list of priority types, such as:

— Critical Features, \( F_C \)
— Important Features, \( F_I \)
Additionally, we propose to annotate all features in a feature group as either required (\(F_R\)) or optional (\(F_O\)) feature group. The priority type and the feature group can be formally defined using the following enumeration:

- Priority List = \(<F_C, F_I, F_U>\)
- Feature Group = \(<F_R, F_O>\)

\(F_C\) always lies in \(F_R\) feature group as service of those features can only be degraded in critical situations but shouldn’t be unavailable until a complete shutdown of the device. On the other hand, \(F_I\) is under control of designer’s justification as it can be programmed as running with degraded service as well as making features unavailable as a last resort to adapt to context changes. For this reason \(F_I\) can stay in both of the feature groups \(F_R\) and \(F_O\). \(F_U\) features are not subjected to feature degradation, those will be unavailable in critical situations. As \(F_U\) is nice to have features, they are in \(F_O\) feature group. They are not subjected to feature degradation instead either the context allows their activation or make them unavailable.

Formally defined, priority-adaptability mapping takes priority type and feature group of each of the features and maps them to the possible adaptability that needs to perform. Table 1 identifies the mapping we are suggesting where the critical feature can only be adapted by degrading their service. Important ‘required’ features should be degraded first but they can be made unavailable in worst cases. Important ‘optional’ features and useful features should not be programmed to provide degraded services instead they should be taken off if needed.

### Table 1: Feature Priority-Adaptability Mapping

<table>
<thead>
<tr>
<th>Input (Feature Prioritization)</th>
<th>Output (Adaptability Mapping)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Priority Type</td>
<td>Feature Group</td>
</tr>
<tr>
<td>(F_C)</td>
<td>(F_R)</td>
</tr>
<tr>
<td>(F_I)</td>
<td>(F_R)</td>
</tr>
<tr>
<td>(F_O)</td>
<td></td>
</tr>
<tr>
<td>(F_U)</td>
<td>(F_O)</td>
</tr>
</tbody>
</table>

#### 3.1.3 Context Injection

In this phase, requirements models are updated to fit the features modeled in Feature Modeling phase. This phase uses the Feature Priority-Adaptability Mapping as an input so that the requirements models can be updated using this mapping. Now-a-days, varieties of requirements models used in requirements engineering phase depending on the nature of the applications. Here, we examined our framework using UML activity diagrams as it is the most common requirement model used in different software projects, though other models can fit here easily. As UML activity diagrams organize the functionalities provided by an application in workflows, the context can be easily inserted into the decision states of the activity diagrams. The proposed framework insist software engineers to follow following steps to inject contexts in requirements models:

1. At first, engineers design separate activity diagrams for each of the features
2. Then, each of the alternative features of an abstract functionality is merged into a single activity diagram
3. In overall activity diagram, different functionalities are enumerated using 1, 2, ... while each of the alternative features are annotated using A, B, C, ...

A prototype of this extended activity diagram is illustrated in Fig. 1. The illustrated activity diagram models three abstract functionalities (1, 2 and 3) and 7 features (1A, 1B, 2A, 2B, 2C, 3A and 3B). First functionality is implemented using two alternative features, whereas second one has three alternative features and third one comprises two alternative features.

#### 3.1.4 Analytical Modeling

In [1], we use Markov Decision Process (MDP) as the analytical model. There are several reasons behind our choice to refer MDP. MDPs are simply conventional state-transition systems with annotations on transitions and software engineers use state-transition systems very frequently in practice while designing and analysing the software systems. In addition, the initial model of MDP can be expressed as an activity diagram or any other control-flow model of the application and there are several approaches available in literature to support automatic model-to-model transformation from activity diagrams to MDP. An example is the ATOP [19] tool.

The transformation process comprises three sequential steps.

1. It first translates activity diagrams of each abstract functionality into simple MDP with an initial state, a final state and as many as intermediate states as the number of alternative features associated with the abstract functionality.
Next, it adds non-deterministic transition from the initial state to each intermediate state, in turn connected with the final state (illustrated in Figure 2). These non-deterministic transitions are labeled with contextual requirements.

Finally, the resulting MDPs are merged into a single MDP that represents the transition of the overall Activity Diagram.

Fig. 2: MDP Translation Process

3.2 Runtime

The overall MDP generated at design time is used to verify the runtime condition of the system against the minimum context requirements. This verification process is done using a stochastic model checker [20] which can check runtime contextual properties against the minimum requirements and generates an alert if it finds any divergence from the ideal requirements. This helps to determine the adaptation decision proactively as the verification process is being performed continuously and alert comes before execution of the feature. For implementing this task properly, we present an adaptation manager which is composed of three modules: Observer, Interpreter and Executor.

3.2.1 Observer. Observer is programmed to collect data regarding to operational contexts. It relies on the hand-held device operating system to collect those information. All popular operating systems for smart phones and tablets offer a set of public APIs for developers to collect sensed information and make use of them. This module should be programmed to run continuously as it can use the APIs offered by the device operating system and collect context information continuously.

3.2.2 Interpreter. Interpreter is responsible for execution of analytical model using stochastic model checker. It continuously navigates through the each state of the analytical model and checks each of the contextual factor against the collected context data. Interpreter executes following calculation and generates an alert if necessary:

—For single-valued contexts, it simply executes a conditional statement that returns true if the context is available and returns false if the context is unavailable.

—For threshold-based contexts, at first it calculates the summation of minimum values of each context that is required to reach the final state. Then, it checks the summation against the available context. Finally, it returns true if available context is greater than the summation of minimum values, otherwise returns false.

Returning a false by interpreter signifies an alert should be generated for the corresponding feature.

3.2.3 Executor. This module is responsible for executing the feature unavailability or degradation mechanism. It switches the status of the features depending on the contextual criteria. For this purpose, we use Bundle Manager Algorithm of [10]. Bundle Manager Algorithm uses threads to manage different states depending on contextual criteria. If ’Bundle Manager’ receives any alert from the interpreter, it starts a new thread with updated ‘available feature set’ and destroys the running thread. To ensure uninterrupted service, it locks the running thread if it finds any feature running. That means, deactivation of an operational feature will be postponed until the operation is completed. By doing this, it maintains the runtime consistency.

Table 2: Alternative Features

<table>
<thead>
<tr>
<th>Functionality</th>
<th>Contextual Factor: Min Requirements</th>
<th>Features</th>
</tr>
</thead>
<tbody>
<tr>
<td>Contact with doctors</td>
<td>Network: Strong Battery: Full</td>
<td>Real time video</td>
</tr>
<tr>
<td></td>
<td>Network: Average Battery: Medium</td>
<td>Real time voice</td>
</tr>
<tr>
<td></td>
<td>Network: Poor Battery: Low</td>
<td>Real time chat</td>
</tr>
<tr>
<td>Locate Nearest Emergency</td>
<td>Network: Average Battery: Medium</td>
<td>Display Map</td>
</tr>
<tr>
<td></td>
<td>Location: GPS</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Network: Low Battery: Low Location:</td>
<td>Display Text</td>
</tr>
<tr>
<td>Receive Prescription</td>
<td>Network: Medium</td>
<td>Download Detail</td>
</tr>
<tr>
<td></td>
<td>Network: Poor</td>
<td>Show Summary</td>
</tr>
</tbody>
</table>
Next, each of the features modeled by generating a template for their concrete implementations.

Listing 2: Feature modeling of Locate nearest hospitals

```java
@Functionality (name = "LocateNearestHospitals")
@Feature (name = "LocateNearestHospitalByMap")
@Level (metrics = {"Network", "Battery", "GPS"}, values = {2,2,1})
public String automaticLocationIdentifier (String position) {
  // invoke GPS service
}
public String displayMap () {
  // invoke Map service
}
@Feature (name = "LocateNearestHospitalByText")
@Level (metrics = {"Network", "Battery", "GPS"}, values = {0,1,0})
public String manualLocationIdentifier () {
  // ask user to insert his/her current location
}
public String displayText (String location) {
  // show address querying database
}
```

4. EXPERIMENT

We design a mobile distributed emergency response system that is intended to aid during medical emergencies. We named it ‘e-Health Care’. Though this application may offer more services in practical, for simplicity we limit those to following ones:

1. Contact with the help center doctors
2. Receive Prescription
3. Locate nearest hospitals

We implemented e-health Care application using two different approach. The first approach follows traditional application development methodology which offers a static set of features. We named it as Non-Adaptive application. In this application all the features were available throughout the life span of the application. The second approach followed the proposed framework which updates available feature set dynamically at runtime depending on contextual factors. The identified alternative feature sets are illustrated in Table 2. An example of the alternative implementations is shown in Fig. 2. The overall activity diagram of e-health Care is illustrated in Fig. 3. Using the activity diagrams, the corresponding MDPs are generated. Firsts, we generated three states for three functionalities annotating F1, F2 and F3. Here, F1 signifies Contact with doctors functionality, F2 is Locate nearest hospitals functionality and F3 signifies Receive prescription functionality. Then, we generated different states for each of the features and augmented an automaton by creating paths from the states of their corresponding abstract functionalities. Each of the path is annotated with the minimum context requirements needed to execute the feature. The overall MDP of e-health Care is illustrated in Fig. 4. During our experiments, the test have been conducted using LG G2 mobile device with Android version 5.1.1.

4.1 Context-Awareness Testing

To evaluate the usability increased by our framework, we examined the battery saving increased by our framework. First, we ran the non-adaptive version of the application. In this case, we attempted to use all the features of the application repeatedly until the complete shutdown of the device. We repeated the test several times starting at different battery level each time.

Fig. 3: Overall Activity Diagram of e-health Care

Fig. 4: Overall MDP of e-health Care
On the second phase, we ran the adaptive version of the application. The test was done by running the test application under same conditions that were used in non-adaptive version. In this case, if the service of the feature is available, we use the feature, otherwise we move to next feature, and so on. The adaptive application has increased on an average 30.46% battery life than non-adaptive application. The result is illustrated in Figure 5.

Figure 5 shows the proposed framework extends the lifetime of the battery more than 3 hours. In other words, user can get a considerable additional time while using any application that implements the proposed framework. A point to be noted here that, until 70% of battery, the adaptive and non-adaptive applications will have the same battery consumption. However, things will change when battery reaches 70% as the adaptive application will move to change its configuration biased towards energy saving.

4.2 Runtime Overhead Testing

Since the framework requires additional runtime computation, we verified the overhead imposed by the framework. First, we measure the battery usage to enable and disable threads. The experiment result is shown in Fig. 6. From Fig. 6 we can see that activity of stopping and activating features had no significant impact on the battery, only 1.5% battery was consumed when we deactivated 7 features at once. In real world, it is a very rare case that 7 features are affected by context changes simultaneously.

We also measure the time to enable and disable the features. The result is illustrated in Fig. 7. Fig. 7 shows that the time needed to activate and stop bundles is in the range of milliseconds and it grows linearly with the number of features. The more features stop and activate at once, the more time needed to achieve this task. Having said that, activating and deactivating 7 features at a time will need an average of 0.6s, which is very insignificant according to user experience.

5. CONCLUSION

This paper identifies the necessity of a light-weight and proactive adaptive framework for mobile applications and proposes an approach that supports design, development and runtime management of the self-adaptation mechanism in mobile applications. It uses the requirements models of the model-driven software development framework to make the adaptation decision proactively. To minimize the complexity of the adaptation strategy, it uses dynamic selection of mutually exclusive features instead of architectural reconfiguration. The outcome of this adaptive framework is really
significant while the overhead imposed by the framework is negligible according to user experience.

6. REFERENCES


