Neurocomputing ∎ (∎∎∎∎) ∎∎∎–∎∎∎

Contents lists available at ScienceDirect

Neurocomputing

journal homepage: www.elsevier.com/locate/neucom

Optimal resource usage in ultra-low-power sensor interfaces through context- and resource-cost-aware machine learning

Steven Lauwereins [a,*], Komail Badami [a], Wannes Meert [b], Marian Verhelst [a]

[a] KU Leuven ESAT-MICAS, Kasteelpark Arenberg 10, Heverlee B-3001, Belgium
[b] KU Leuven CS-DTAI, Celestijnenlaan 200A, Heverlee B-3001, Belgium

ARTICLE INFO

Article history:
Received 30 June 2014
Received in revised form
23 October 2014
Accepted 17 November 2014

Keywords:
Resource cost-aware classifier
Context-aware machine learning
Low-power sensor interface
Adaptive circuits
Power restricted classifier

ABSTRACT

This paper introduces an approach that combines machine learning and adaptive hardware to improve the efficiency of ultra-low-power sensor interfaces. Adaptive feature extraction circuits are assisted by hardware embedded training to dynamically activate only the most relevant features. This selection is done in a context- and power cost-aware manner, through modification of the C4.5 algorithm. As proof-of-principle, a Voice Activity Detector illustrates the context-dependent relevance of features, demonstrating average circuit power savings of 70%, without accuracy loss. The RECAS database developed for experimenting with this context- and dynamic resource-cost-aware training is presented and made open-source for the research community.

1. Introduction

Interest in ubiquitous sensor networks is strongly increasing, spearheading applications relying on smart objects and smart environments. The sensors in these networks are expected to operate autonomously and continuously throughout their complete lifetime of multiple years. This restricts the average power budget of such sensors to a few μW, which is difficult to attain for current sensors [1,2]. It is therefore important to discard irrelevant data as early as possible, in order to not waste scarce resources on the incoming sea of data. It has, for example, been well established that discarding irrelevant data before wireless transmission by on-board processing is far more power cost-effective compared to transmitting the raw data to a central data collecting node [3]. Taking this one step further, also within a sensor node, relevant features should be extracted as close to the raw sensor as possible to avoid power wastage.

At the same time, there is a trend within the hardware community to reduce the power consumption of their designs through the development of dynamically scalable and adaptive hardware [4–7]. This kind of hardware can be reconfigured to the most power efficient configuration while meeting the requirements of the current use case. This use case is on average less stringent than the most demanding case, thus allowing for reduced power consumption. The variable behavior of the operational conditions and use cases is too complex and unpredictable to apprehend at design-time and is therefore best handled by autonomously exploiting this adaptivity. However, autonomously and optimally configuring these adaptive systems is still difficult. Current state-of-the-art (SoA) dynamic self-reconfigurable systems focus on optimizing power hungry hardware blocks such as DSPs for video compression and multi-core processors [8,9]. Straightforward application of this dynamic self-reconfigurability paradigm is however impossible for power-scarce sensor nodes, due to the significant power overhead of the required embedded adaptivity management.

In previous work [10–12], we introduced and proved on silicon, a new sensing paradigm where hardware self-adaptivity is exploited within power-scarce sensor interfaces. Relevant information (features) are extracted as close to the raw sensor as possible to be power-efficient. The targeted sensor interfaces selectively extract features such that hardware resources of unselected features can be turned off to save power. This operating paradigm promises significant power savings (up to 10 ×) [10–12], however its implementation requires an advancement beyond the state-of-the-art of hardware self-adaptivity on three fronts.

Firstly, the dynamic control of feature selection and resulting hardware scalability needs to jointly take into account the impact of the scalability on the sensor's power budget, as well as on the sensor's classification accuracy. This results in a novel trade-off between power consumption and classifier accuracy, which can be exploited dynamically.

Secondly, the reconfigurability management has to be implemented with significantly reduced power overhead compared to the current SoA self-adaptive systems.

* Corresponding author.
E-mail address: steven.lauwereins@esat.kuleuven.be (S. Lauwereins).

http://dx.doi.org/10.1016/j.neucom.2014.11.077
0925-2312/

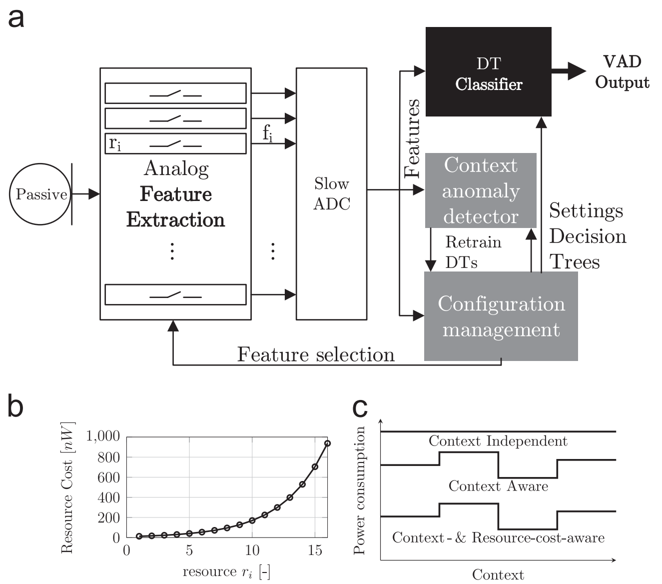

Fig. 1. (a) Context- and feature cost-aware classifier showing integration of adaptive hardware and machine learning. The grey blocks are only sporadically activated. (b) Illustrates the varying power cost across feature extraction hardware blocks. (c) Show anticipated power consumption versus context for three types of classifiers. (All applicable for acoustic sensing scenario introduced in Section 5.)

Thirdly, the configuration of the hardware needs to be optimal at every moment in time. However, environmental changes (from now on called context-switches) can lead to sub-optimal configuration of the overall system. This results in a need for at run-time detection of context-switches and a need for at run-time unsupervised retraining of the optimal configuration.

The here envisioned sensor interfaces overcome those challenges through the introduction of an adaptive feature extraction block dynamically configured with a low-cost machine learning scheme working together with a sporadically activated context anomaly detector (Fig. 1a). The analog feature extraction block allows dynamic deactivation of irrelevant features reducing the power consumption of this block. An additional power reduction is obtained by extracting the features in the analog domain because this requires a lower sampling frequency of the analog-digital-converter compared to digital feature extraction [12,13]. A digital classifier uses the selected features for sensor signal classification with a target accuracy constraint. Embedded machine learning hardware is responsible for the optimal configuration of this feature extraction block and the definition of the classifier at any given moment in time. It therefore implements an unsupervised configuration management unit as well as a context anomaly detector which triggers retraining, Fig. 1a.

To reduce the overhead of this adaptivity management and the embedded classifier, the classifiers are realized using decision trees. Such trees enable power-efficient inference through the implementation of a decision tree classifier in dedicated hardware. Moreover, the embedded training of the decision tree inherently reveals the current importance of features, which is exploited to facilitate the dynamic feature selection process. Our previous work [10–12] showed that such low-cost machine learning is capable of controlling adaptive sensor node hardware resources within the sensor node's energy budget. This paper extends our previous work along four axes:

1. Presenting a detailed description of the algorithmic modifications to the C4.5 algorithm [14] made for the extension to true resource-cost-awareness.
2. Enabling optimal unsupervised training of a classifier under a predefined power budget through dynamic tracking of resource-selection within a best-first tree training scheme.
3. Enabling context-switch detection through one-class decision trees.
4. Release of our open-source context- and resource-cost-aware feature database (RECAS) for a voice activity detector, made available to the research community.

In this regard, Section 2 introduces the notion of context-aware feature activation, to dynamically activate current most relevant features. Section 3 further expands this idea to context-aware and dynamic resource-cost-aware classification for improved power efficiency, taking also resource circuit power cost into account. It discusses in depth the necessary algorithmic changes required for optimal training. Section 4 describes the method used to train decision trees in an unsupervised way and the method developed to detect context-switches. Section 5 applies the derived approach on a proof-of-principle hardware design of a Voice Activity Detector, demonstrating up to $10\times$ reduction in power consumption or up to 10% increase in accuracy. Section 5.1 introduces our RECAS database allowing context- and dynamic resource-cost-aware simulations, made available to the research community in an open source scheme.

2. Context-aware feature selection

In many applications, the relative information content of a feature is highly context-dependent. Depending on the context, some features port a more distinctive value towards the classes of interest. Examples are acoustic classifiers, prone to various types of background noises, or patient-specific biomedical data classification [15].

To always operate the target sensing systems at maximal resource efficiency, feature selection should be done in a context-aware way. This context-aware feature selection allows for two different approaches. Firstly, by only activating discriminative features within the current operating context, the amount of extracted features and active hardware resources is dynamically adapted (see Fig. 1a). This reduces the overall power consumption (Fig. 1c) without losing classification accuracy. Alternatively, by activating all features, the context specific training of a tree makes it possible to increase the accuracy due to a tighter cluster of data points within each class.

Both approaches are illustrated in Fig. 2, which shows the ROC curves of a voice activity detection classifier trained and tested on signals with babble background noise. Fig. 2a shows the classification accuracy for a context-aware (CA) classifier with feature deactivation and for a context-unaware (CU) classifier. As previously mentioned the overall power consumption decreases, in this case with a factor 2, without losing classification accuracy. Fig. 2b shows the classification accuracy for a context-aware (CA) classifier trained on the correct context (solid line), for a context-aware classifier trained on the wrong context (dotted line) and for a classifier trained on all contexts, i.e. a context-unaware (CU) classifier (dashed line). As can be expected, accuracy of a classifier trained on a correct context-restricted dataset (e.g. babble noise) significantly outperforms the context-unaware classifier. A wrongly trained classifier, i.e. a classifier trained on the wrong context, performs poorly. This observation introduces a new trade-off when training the model, classification accuracy can be exchanged for lower power consumption.

In this work, decision tree based classifiers will be trained for different contexts. Within every context, the features used in the different nodes of the tree determine the discriminative features requiring activation. As will be demonstrated in Section 5, this context-awareness allows cutting the number of actively observed

Please cite this article as: S. Lauwereins, et al., Optimal resource usage in ultra-low-power sensor interfaces through context- and resource-cost-aware machine learning, Neurocomputing (2015), http://dx.doi.org/10.1016/j.neucom.2014.11.077