Übungen zur Vorlesung

Algorithms for context prediction in ubiquitous systems

Wintersemester 08/09 Stephan Sigg (sigg@ibr.cs.tu-bs.de) Exercise: 11.11.2008, 13:15 – 14:45

1 Context prediction

1.1 High-level and low-level context prediction

In the lecture we have discussed the possibility to implement context prediction either on high-level or on low-level context time series. The approximate error probabilities are

$$P^{approx}_{low-level}(i) = P^{km}_{acq} P^{m}_{pre}(i) P^{o}_{int} \tag{1}$$

$$P^{approx}_{high-level}(i) = P^{km}_{acq} P^{o}_{pre}(i) P^{ko}_{int}. \tag{2}$$

These formulae are derived as follows. For high-level context prediction, the context acquisition step is the first processing step applied to the sampled contexts in form of raw data. For all k time series elements in the context history, every one of the m raw data values is transformed to low-level contexts in the context acquisition layer. Since P_{acq} describes the probability that no error occurs in one of these context acquisition steps, the probability that no error occurs in any of the $k \cdot m$ context acquisition steps is P^{km}_{acq} consequently.

In the context interpretation layer, the m low-level contexts of every one of the k context time series elements in the low-level context history are interpreted to o high-level contexts that constitute a time series element of the high-level context time series. Altogether, $k \cdot o$ context interpretation steps are applied in the interpretation layer. Since P_{int} describes the probability that no error occurs in one of these interpretation steps, the probability that no error occurs in the whole context interpretation process is consequently P^{ko}_{int}. Finally, $P_{pre}(i)$ describes the probability that the prediction of the i-th context is without error. Since the i-th time series element consists of o context elements (o context elements share the same timestamp), $P^{o}_{pre}(i)$ is the probability that no error occurs in the context prediction step. Together, with probability

$$P^{approx}_{hl} = P^{km}_{acq} P^{ko}_{int} P^{o}_{pre}(i) \tag{3}$$

no error occurs in any of the context processing steps utilised for the prediction of one specific high-level time series element. The term P^{approx}_{hl} is calculated accordingly.

a) Can you see from these formulae if high-level or low-level context prediction has the higher error probability?

1.2 Exact calculation

In the calculation in exercise 1.1 we did not take into account that an error in the context interpretation step might correct an error that occurred in the context acquisition step, or that a context prediction error has a correcting influence on erroneous high-level contexts. The probability P^{int}_{cor}

that an error which occurs in the context acquisition step is corrected by an error that occurs in the context interpretation step is

$$P_{cor}^{int} = (1 - P_{acq}^{m})(1 - P_{int}^{o})\frac{1}{v_h^o - 1}. \quad (4)$$

In this formula, $1 - P_{acq}^{m}$ is the probability that an error occurs in one of the m context acquisition steps that are related to one context time series element and $1 - P_{int}^{o}$ describes the probability that an error occurs in one of the o context interpretation steps. However, in this case, no arbitrary error is required but the one interpretation error that leads to the correct high-level context. Since v_h high-level contexts are possible for every one of the o high-level contexts in one time series element, the number of possible high-level time series elements is v_h^o. Consequently, the number of possible errors is $v_h^o - 1$ since one element represents the correct interpretation that is without error. With probability $\frac{1}{v_h^o - 1}$ the required specific error required for a correction is observed out of all $v_h^o - 1$ equally probable interpretation errors.

We now consider the probable correcting influence of the context prediction error. Since we have assumed that every one of the $v_h^o - 1$ incorrect time series elements is equally probable for any incorrectly predicted position i in a predicted time series, the probability, that the correct time series element at position i is predicted from an erroneous context history is $\frac{1}{v_h^o - 1}$. Altogether, the probability $P_{hl}(i)$ that time series element i is accurately predicted if the prediction is based on the high-level context time series is therefore

$$\begin{aligned} P_{hl}(i) &= \left(P_{acq}^{m} P_{int}^{o} + P_{cor}^{int}\right)^k P_{pre}^{o}(i) \\ &+ \left(1 - \left(P_{acq}^{m} P_{int}^{o} + P_{cor}^{int}\right)^k\right) \frac{1 - P_{pre}^{o}(i)}{v_h^o - 1}. \end{aligned} \quad (5)$$

Note that we consider interpretation and acquisition errors separately for every one time series element. This is expressed by the exponent k which affects the term $P_{acq}^{m} P_{int}^{o} + P_{cor}^{int}$ as a whole.

In analogy to this discussion, we receive the probability that a predicted high-level time series T of length $|T| = n$ contains no inaccurate context time series element as

$$\begin{aligned} P_{hl} &= \left(P_{acq}^{m} P_{int}^{o} + P_{cor}^{int}\right)^k P_{pre}^{o} \\ &+ \left(1 - \left(P_{acq}^{m} P_{int}^{o} + P_{cor}^{int}\right)^k\right) \left(\frac{1 - P_{pre}}{v_h^n - 1}\right)^o. \end{aligned} \quad (6)$$

In this formula, P_{pre} depicts the probability that a one-dimensional time series of length n is correctly predicted. Since the dimension of the predicted time series is o, P_{pre}^{o} describes the probability that this o-dimensional time series is error free. The term $\left(\frac{1 - P_{pre}}{v_h^n - 1}\right)^o$ depicts the probability that an error in the interpreted time series that occurs with probability

$$\left(1 - \left(P_{acq}^{m} P_{int}^{o} + P_{cor}^{int}\right)^k\right) \quad (7)$$

is corrected in the prediction step. Again the prediction errors are considered separately for every dimension of the high-level context time series, since the exponent o affects the term $\frac{1 - P_{pre}}{v_h^n - 1}$ as a whole.

a) Calculate the corresponding probability for context prediction based on low-level context time series.

1.3 Determination of the predominant approach

How would you demonstrate for these exact probabilities which prediction approach (high-level or low-level) has the higher probability of error in a given scenario?