DOI 10.5162/SMSI2023/A4.2

Correction of Inconsistent Sensor Timing: Missing Samples and Clock Deviations

Yogesh Kapoor[1], Walter Lang[1], Reiner Jedermann[1]
[1] Institute for Microsensors, -Actuators and –Systems (IMSAS),
University Bremen, Otto-Hahn-Allee 1, Germany
rjedermann@imsas.uni-bremen.de

Summary:
Sensor data, transferred by wireless sensor networks, must often be resampled for correcting missing samples or timing deviations. The direct application of the well-known Shannon theorem is limited to uniform sampling. For the non-uniform case, the resampling accuracy can be improved by Local Regression. Several cases, such as low/high signal frequency and noise were tested in combination with different scales of non-uniformity from missing single samples to fully arbitrary sampling.

Keywords: Arbitrary resampling, missing samples, clock jitter, Whittaker-Shannon Interpolation, sensor data processing, non-uniform sampling.

Problem and Motivation

The timing of measurement data, transmitted via wireless sensor networks, is often inconsistent. Samples might be missing or delayed due to network failures or overload. The CPU clock of the sensor nodes might deviate or be out-of-sync with other sensor nodes. Subsequent processing of the sensor data mostly requires uniform sampling at identical time points for different sensor nodes. The measurements must be resampled accordingly for digital twins and IoT based data processing.

In practice, often only the simple approach of "keeping the last measured value" is applied, or, in technical terms, a Zero Order Hold function (ZOH). The Whittaker-Shannon Interpolation (WSI) according to the sampling theorem of Shannon [1] offers full reconstruction from a theoretical point of view under some restrictions by filtering the input with a sin(x)/x function. The first restriction, that the signal is bandlimited to half of the sampling rate, is mostly assumed to be fulfilled. However, other restrictions often cannot be complied with in practice. Sampling must be uniform and free of noise. The signal must be infinite in time. Especially the latter one is problematic for real-time processing of sensor data, where only the past values are known.

In this contribution, we test different alternate methods, in order to provide an improved solution adapted to the sampling conditions.

Methods and test cases

In addition to WSI and ZOH, we tested Local Polynomial Regression (LPR) [2], also known as Locally Estimated Scatterplot Smoothing (LOESS), and linear connection of neighboring measurement points (LIN).

Sine waves at different frequencies, a random signal created by Gaussian process, and temperature measurements in a building were applied as test signals. Reference signals were generated at 10fold sampling frequency. A set of samples was picked from this reference set at lower sampling frequency as example measurement data. Signals with low/high frequency and with/without noise were tested. Besides uniform sampling for f=1Hz, different inconsistent timing conditions were tested, such as missing samples, clock jitter, and arbitrary time points with a maximum distance of 1.5 s.

For comparing the accuracy of the different methods, they were applied to resample the signal to the reference frequency. The Root Mean Square Error (RMSE) was calculated between the resampled and the reference signal. The reconstruction was tested for the offline case with a measurement set of defined length, as well as real-time case where only past measurements are known, and the reconstructed signal must be updated after each new measurement.

Results

Fig. 1 shows a test signal created by a Gaussian process with a covariance width of 1. Different algorithms were tested to reconstruct the full signal by samples takes in intervals of 1 s. All methods had problems to approximate short peaks in the signal. Best results were achieved

DOI 10.5162/SMSI2023/A4.2

by WSI with an RMSE of 0.060 followed by LPR with 0.111.

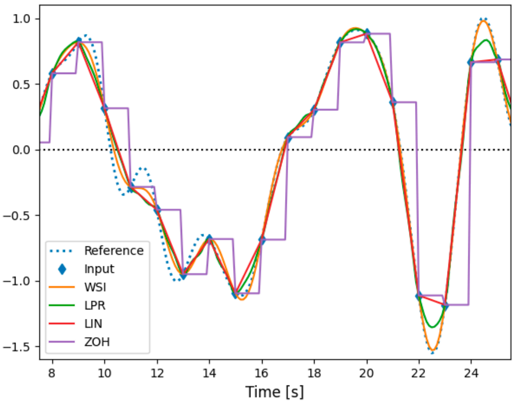

Fig 1. Gauss process test signal (covariance = 1) and resampling

For further testing of the frequency sensitivity of the different methods, sinus waves were applied as test signal. WSI is the method of choice for signals with frequencies close to the theoretical limit for the relation between highest signal and sampling frequency of r=50% (Fig. 2). If r drops below 20%, LPR provides the same or even better accuracy. For low frequencies r<3%, LIN can be used without losing accuracy.

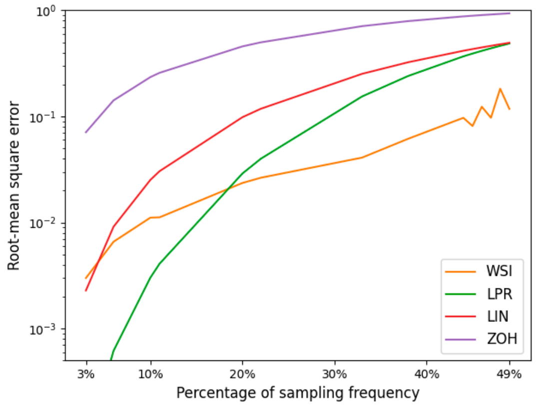

Fig. 2. Prediction error as function of frequency

The effect of inconsistent sensor timing was further tested for WSI and LPR. For the test of clock jitter, the sampling time point was varied by ±10% of the sampling interval. In a second test, an updated signal prediction was calculated after each new measurement with only the past data known.

WSI turned out to be very sensitive towards deviating clock timing, or cutting the input to past values, as the high RMSE indicates (Fig. 3, orange solid lines).

LPR (green) was hardly affected by jitter but for the update case, the RMSE increased by a factor ~3. Leaving out two samples gave about the same error as the update case. Arbitrary sampling gave similar results as the update case for LPR, but the highest RMSE for WSI.

For signal frequencies below 20% of the sampling frequency, LPR provides acceptable accuracy for non-uniform sampling, but for higher frequencies, an accurate method is still missing.

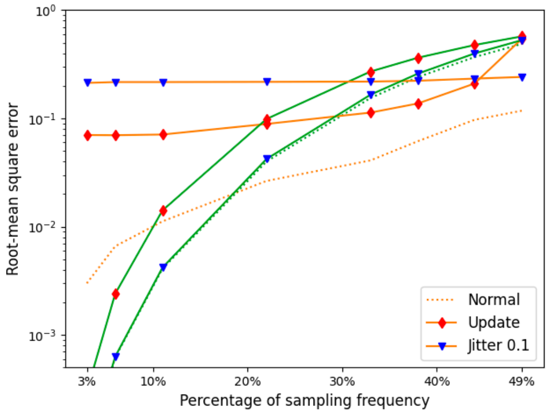

Fig 3. Effect of non-uniform sampling (WSI orange, LPR green)

Discussion

LIN was always better than ZOH for all test signal frequencies, e.g., by a factor of 4 for r=20%. The error can be further reduced by using LPR, e.g., by an additional factor of 3 at r=20%. Only for higher frequencies with r>20%, WSI achieved better results.

LPR is also suitable for the update case and non-uniform sampling with r<20%. However, there is still a lack of good methods for higher frequencies with r>20%. Last results concerning improved methods will be presented at the conference, including Kriging [2], Akima [3], and a newly developed modified local regression method. A Digital Twin platform for sensor data processing was presented at the previous conference [4]. The suggested resampling methods provide the necessary extension for inclusion of sensor data with timing deviations.

References

[1] J. M. Whittaker, Interpolatory function theory: The University Press (1935)

[2] R. Jedermann, Feasibility of Low Latency, Single-Sample Delay Resampling - A New Kriging Based Method, *Algorithms* 16, no. 4: 203; doi: 10.3390/a16040203

[3] A. Patki, G. Thiagarajan, Low complexity, low latency resampling of asynchronously sampled signals, *International Conference on Signal Processing and Communications (SPCOM)*, 2016

[4] R. Jedermann, K. Singh, W. Lang, P. Mahajan, Digital twin concepts for linking live sensor data with real-time models, *J. Sens. Sens. Syst.*, 2023, Vol. 12(1), pp. 111-121; doi: 10.5194/jsss-12-111-2023