Data Pre-processing on Sensor Nodes for Predictive Maintenance

Aleksandr Timofeev, CEO
Sensor technologies are experiencing exponential growth with forecasts of ~45 trillion sensors in 2032 that will generate >1 million zettabytes ($10^{27}$ bytes) of data per year (1).

Sensor raw data processing is a bottleneck.

Use of deterministic operations is a barrier for energy efficiency.

True parallel data processing is the key for unprecedented energy efficiency.

Biological systems are 100,000 times more energy efficient (2) than digital computer systems.

(1) https://www.frontiersin.org/research-topics/31093/nanoelectronic-devices-for-analog-hardware

(2) Landauer's principle defines the theoretical limit of energy consumption of computation.
SOLUTION: NEUROMORPHIC ANALOG SIGNAL PROCESSING OF SENSOR RAW DATA

- Analog neuromorphic cores based on true parallel process extract useful information from sensor raw data more efficiently than traditional digital methods based on deterministic operations.

- Extracting useful information from sensor raw data is based on mimicking biological sensing processes of nature.

![Diagram showing the process of sensor data processing through conventional and neuromorphic analog methods.](image)

- An analog neuromorphic front-end with a neural network implemented in analog on the chip is several orders of magnitude more efficient than a digital system deterministic calculations of a math model.
Neuromorphic Analog Signal Processing technology platform is the key to developing and manufacturing the neuromorphic front end.

NASP technology allows conversion of trained neural networks into an analog neuromorphic core and its manufacturing.

Select Trained NN

- Trained Neural Net math model in any standard format: Keras, TensorFlow or others
- Assisted and supported by POLYN

Convert Math Model Into Chip Layout

- Math model generated with NASP Compiler
- D-MVP – Product digital simulation for validation
- Chip model Netlist ready
- Fully automated process by POLYN

Chip Production at a Fab

- NASP GDSII files ready for any foundry: TSMC, Global Foundries, Samsung, etc.
- NASP chips manufactured from standard analog components according to NASP neuron design
- Fully automated process by POLYN
- Any Fab with 40-90 nm process

Chips for the Sensor Node Product
APPLICATION: VIBRATION MONITORING

Example: embedded Temperature Pressure Monitoring (TPM) Sensor with integrated 3-axis accelerometer and VibroSense chip. Proof-of-Concept (PoC) with a tire manufacturer.

NFE enables data volume reduction by more than 1000 times, and power consumption reduction by 100 times to support ultra-narrow band communication and very long battery life or energy harvesting based designs.

TODAY: No data reduction – all traffic to the cloud

With POLYN: data pre-processing on sensor

Ultra low bandwidth
Energy harvesting
Input signal from 1 axis accelerometer >0.6 Mbps for 10 kHz bandwidth

- Data compression by a neural network of efficient size
- Avoid lossy compression
- Necessary to find an encoder function to minimize the information loss or maximize mutual information:

\[
\hat{f} = \arg\max_{f \in F} I(X; f(X))
\]

where \(X\) is an input signal, \(f\) is an encoder, \(I\) is mutual information.

The above is hard to optimize directly. We can minimize the autoencoder loss by searching for a decoder function \(g\) as well:

\[
\hat{f}, \hat{g} = \arg\min_{f \in F, g \in G} L(X, g(f(X)))
\]

Visualization of the encoding process: transforming input signal into latent representation

\[X \xrightarrow{f} Z = f(X)\]
OVERVIEW OF THE NEURAL NETWORK MODEL ARCHITECTURE

WHY WAVELETS:
- better time-frequency resolution balance than FFTs
- good fit to the chip: short shared convolutions

AUTOENCODER VARIANTS:
- **Unrestricted**: highest compression rate but more entangled latent space
- **Latent space regularization (VAE)**: lower compression rate but 'nicer' latent space

Level 0
- Wavelet filter
- Wavelet filter
- Wavelet filter
- Wavelet filter
- CNN₀

Level 1
- Wavelet filter
- Wavelet filter
- Wavelet filter
- CNN₀

Level k
- Wavelet filter
- Wavelet filter
- Wavelet filter
- CNNₖ
METRICS AND RESULTS OF TEST DATA SET

- 38 classes (balanced)
- 11 hours
- 921 millions of data points

Confusion Matrix. Accuracy: 0.942

PCA projection

Well-structured low-dimensional space

T-SNE projection
1 axis, 24 kHz sampling rate, 24 bit ADC

Min 562.5 Kbit/s
**NFE – ENABLING SOLUTION**

**TPM sensor node low-rate data flow**
- Temperature sensor
- Pressure sensor
- 0.032 kbit/s
- 8-bit low power MCU
- CR2405 battery - 5Y lifetime

**TPM + Accelerometer**
- Temperature sensor
- Pressure sensor
- 12 kHz 2-axis Accelerometer
- 0.032 kbit/s
- 8-bit low power MCU
- CR2405 ~ 2 hours lifetime

**At sensor pre-processing with Digital MCU**
- Temperature sensor
- Pressure sensor
- 0.032 kbit/s
- 32-bit ARM Cortex M0+
- >1000 Kbit/s
- CR2405 battery ~ 30 days lifetime

**At sensor pre-processing with NASP VibroSense NFE**
- Temperature sensor
- Pressure sensor
- 12 kHz 2-axis Accelerometer
- 0.032 kbit/s
- 8-bit low power MCU
- >1000 Kbit/s
- Power consumption: <100 µW
- Enabling energy harvesting solutions

**Not possible to transmit such high-rate data flow**

433 MHz or BLE

---

---
VibroSense™ CHIP ARCHITECTURE

Light version – Host control

Disabled mode: Less than **10 nW**
Enabling mode: Less than **100 µW**
Configuration and control - external Host MCU
Seamless integration with existing sensor node designs, such as standard TPMU sensors with integrated 8-bit MCU

Full version – Integrated 8-bit ultra low power MCU

Disabled mode with RTC: Less than **500 nW**
Enabling mode: Less than **1 mW**
Capability to function as a host for LoRa or other RF modules
Ideally suited for industrial IoT modules
NFE is a new technology based on Neuromorphic Analog computation on-chip that can be applied to various tasks.

We invite anyone with a model for a sensor: NASP will convert it into a chip, efficient, sustainable for next generation of IoT devices.
Copyright Notice

This presentation in this publication was presented as a tinyML® EMEA Innovation Forum. The content reflects the opinion of the author(s) and their respective companies. The inclusion of presentations in this publication does not constitute an endorsement by tinyML Foundation or the sponsors.

There is no copyright protection claimed by this publication. However, each presentation is the work of the authors and their respective companies and may contain copyrighted material. As such, it is strongly encouraged that any use reflect proper acknowledgement to the appropriate source. Any questions regarding the use of any materials presented should be directed to the author(s) or their companies.

tinyML is a registered trademark of the tinyML Foundation.

www.tinyml.org