Prototype Implementation of Dynamic Data Pruning in Smart Energy Meter

Payali Das[†], Sushmita Ghosh[*], Wadood Ahmad Khan[‡], Shouri Chatterjee[†], Swades De[†]
[†]Department of Electrical Engineering, [*]UQ-IITD Academy
[†,*]Indian Institute of Technology Delhi, New Delhi, India
[‡]Center for Development of Telematics, New Delhi, India

***Abstract*—Smart electricity meter is an example of the Internet of Things (IoT) device which is nowadays being installed in domestic and industrial premises to monitor the power consumption along with various other services such as billing, load forecasting, and dynamic pricing. Usually, these IoT devices sample at a high rate and report the whole data to the cloud. In this process a large amount of energy is consumed at the IoT node for transmission of the data and a large bandwidth is used for transmission over the Internet, though for storage saving some compression is performed at the destination. To this end, we aim to prune the IoT data at the source node by providing some intelligence, thereby aiding to the wireless IoT node's energy sustainability and efficient channel bandwidth usage for scalable deployment. The IoT devices measuring/reporting single parameter or multiple parameters are judiciously pruned within an acceptable reconstruction error limit. In this paper, we report embedded implementation of data-driven dynamic pruning of multi-parameter smart meter data as an example demonstration of data-smart IoT nodes. Our performance results show that the energy and bandwidth savings with multiple univariate data pruning are respectively about 19% and 36%, whereas the savings with multivariate data pruning are respectively about 36% and 98%. The developed embedded data pruning module is 99.09% more energy efficient than the implementation on Raspberry Pi.**

***Index Terms*—Advanced metering infrastructure, data-driven smart IoT, energy efficiency, prototype design and implementation, smart meter**

I. Introduction

An electric power grid is an interconnection of power generation sources, electric transmission network, transformers, and distribution network to deliver power to the consumers. The operation of conventional electric grid has become obsolete due to the lack of automated analysis; it is unsuitable for meeting the fast-growing and dynamic demand for electricity. Hence, the concept of Smart Grid (SG) has gained prominence which is touted as the next generation of electric power system. SG is considered for interconnecting the various components of electric grids using communication networks [1]. It aims to provide electricity supply to the end users in a safe, reliable, efficient, and secure way. In SG, two-way communication is established between the various components and it allows the end user to not only consume power but also supply excess power to the SG. This is achieved with the the help of smart meters that monitor and measure the bi-directional flows.

Smart Meter provides electricity consumption or supply data to the Head End System on routine basis as well as on demand [2]. Demand side management and load limiting are two key features aided by the smart meter that enables the utility provider to estimate the trend on energy usage of a particular customer or entity [3].

Advanced metering infrastructure (AMI) is a collective term to describe the whole infrastructure that collects and analyzes in near real-time data from smart meters and control center equipment using two-way communication. This data is used to manage various applications and services intelligently. A smart meter is capable of measuring power consumption at a fine granularity as compared to a conventional meter. It periodically sends the collected information to the utility provider for billing, monitoring of load, and demand-response mechanisms by the control center. It also enables consumers to monitor and control their power consumption. Hence, through customer participation, it may be possible for the utility provider to provide electricity at dynamic and relatively low prices. Smart meter is also capable of providing various instantaneous information on circumstances, such as power theft detection and tampering, thus offering to secure the utility [4].

A. Motivation

The huge amount of data generated by the smart metering framework helps in understanding and modeling the patterns of energy usage. However, resource-efficient transmission and storage of this vast data remain a challenging task. Due to limitations in the handling of this vast amount of data, methodologies for smart meter data reduction need to be exploited. While the smart energy meters have been introduced in SG with the view of remote and autonomous two-way data connectivity, it does not have node-level data handling smartness. In other words, data-driven resource optimization techniques have not been exploited in smart meters. Recent research studies indicate that the incorporation of data-level smartness at the smart energy meters can significantly reduce the communication bandwidth requirements [5], [6]. Yet, these studies did not delve into the implementation details and the associated challenges. Moreover, they did not address the energy footprint aspects from the additional node-level processing and reduced communication requirements.

B. Key contributions

In this work, we have developed a cost-effective end-to-end smart energy metering solution prototype with a data pruning subsystem at the source node. The data pruning subsystem employs dynamic data-driven resource optimization techniques depending on the allowable level of compression for a given error toleration. As a use case study, a smart energy meter has been chosen to demonstrate the concept as it makes available multiple parameters that can be used to achieve high levels of compression. This concept can be easily extended to similar IoT devices generating time-series data consisting of single or multiple parameters. To make the prototypes commercially viable, our designed smart data handling and communication module has the following features:

- Reduced implementation complexity: Space complexity of the algorithm arises since the system needs to work in near real-time, extra memory will be required to store the intermediate values generated during the operation of the algorithm. So, optimized implementation of the algorithm will play a crucial role in reducing the cost as it will have a small memory footprint.
- End-to-end solution for commercial deployment: The system should support popular communication technologies. It should also be upgradable to support technologies to be deployed in the near future. This will tend to increase the cost of the product. This should be made flexible in order to cater to different markets.
- Energy minimization: By designed energy efficient smart IoT node consumes 19.7% and 36% less energy during transmission by incorporating univariate and multivariate data pruning algorithms respectively.
- Cost minimization: The choice of the hardware components were in such a way that we can achieve a reasonable performance at a minimal cost. The designed smart IoT node is 3 times cost-effective than the conventional Raspberry Pi (R-Pi) based implementation.

II. Background and state-of-the-art

Smart IoT communication [7] proposes a data-driven approach. In this framework, each IoT device has been imparted some intelligence to fulfill a certain objective. One approach of doing so is to impart the capability to the edge IoT device to understand the dynamics of the underlying process to convert bulk data into smart data. Imparting data pruning intelligence at the edge node level has two benefits, these are reduced traffic in the communication network and reduced the data storage requirements at the cloud.

In [8], proposes a technique to improve the energy sustainability of the edge IoT node by incorporating data pruning intelligence. Data-driven resource optimization techniques can be employed in commercial smart meter to make them intelligent in the context of data bandwidth saving. Paper [9] suggests that by learning the channel variability and using a dynamic channel coding scheme, the transmission reliability of IoT data can be improved. Implementation of metering systems with various communication technologies are described in [10], [11]. In [10] , a LPC2148 microcontroller incorporating GSM technology based smart home controller is proposed to do net metering and smart appliance control. A single phase energy meter is described in [11] where the communication medium is WiFi. However, none of the works have incorporated any data compression algorithm to reduce the data or memory footprint.

Currently, deployed commercial IoT devices lack the intelligence to exploit data reduction techniques at the source node; it transmits all the sensed data, thereby sending redundant information content. These IoT devices can be considered as conventional sensing devices with network connectivity. With the advent of 5G, there is expected to be a very dense deployment of IoT devices, and the data sampling rates are going to increase [12]. Hence, the need arises to make these devices smarter in some context. To the best of our knowledge, currently there are no such subsystems available with commercial smart meter which can exploit data bandwidth saving strategies.

III. Data Driven Resource Optimization Approaches

Commercial smart meters sense various energy consumption parameters such as current, voltage, apparent power, frequency, etc. along with meter health-related parameters. Unlike conventional analog meters, these automated meters follow a high-speed data sampling approach, thereby generating a huge chunk of data. Data compression approaches employed in smart meters can be classified based on whether they operate at the IoT device or the aggregator in the AMI. Techniques employed at the aggregation point have a high compression ratio since the aggregation point has access to vast amounts of data from many meters connected to it. The sampling rate is typically half-hour, and hence it aids in identifying various energy consumption patterns such as daily, weekly, seasonal, or behavioral. Various compression techniques have been suggested in literature and can be studied under two heads: lossless compression techniques and lossy compression techniques. The lossless compression techniques lead to an accurate reconstruction of data whereas, the lossy techniques report a higher compression ratio and are suitable for error-tolerant scenarios. Our study reveals that a data pruning subsystem employing these algorithms has not been integrated into commercial smart meters.

A. Non-real-time univariate approach

The work in [5] proposes a lossy data compression technique based on compressive sampling on a single time-series data. Compressive Sensing makes use of the sparsity in the given data or time-series to compress it and also provides fairly accurate reconstruction from fewer samples as compared to that required by Nyquist sampling theorem. Consider a data transmission window of n number of samples, where the i^{th} sample is denoted by x_i. Hence, $f = (x_1, x_2, \ldots, x_n)$ are the samples comprising x. So x can be represented as:

$$\mathbf{x} = \psi \mathbf{f}, \tag{1}$$

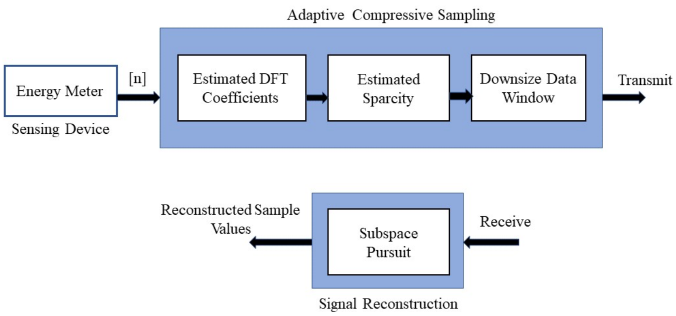


Fig. 1: Flow diagram of univariate approach.

where ψ is a matrix of size $n \times n$ representing the sparse basis matrix and the column vector of coefficients corresponding to ψ are denoted by f. Out of n samples in a given batch of data, only m ($m << n$) samples are randomly chosen for transmission. In order to save transmission bandwidth over the communication channel, this data downsizing is performed. Using 1, the samples which are transmitted can be represented by,

$$\mathbf{y} = \phi\mathbf{x} = \phi\psi\mathbf{x}, \tag{2}$$

where ϕ is sensing matrix of size $m \times n$. To accurately reconstruct the signal using the received samples, the problem of an underdetermined system of linear equations needs to be solved. Subspace pursuit algorithm has been employed in this work to reconstruct the original signal at the cloud server. For signal reconstruction, Random Gaussian Matrix and Discrete Fourier Transform (DFT) are chosen as sensing matrix and sparse basis, respectively, such that restricted isometry and incoherence property are satisfied. In this scheme, the sparsity is computed for each batch of data by calculating the number of DFT coefficients comprising 99.99% energy of samples. This helps in capturing the erratic nature of smart meter data and also plays an important role in compressing the data by decreasing the number of samples which are transmitted without comprising the information content or the reconstruction accuracy. Hence, this scheme is adaptive as compared to the conventional technique of compressive sampling, where the sparsity is inferred and remains the same. Accordingly, m= s log n are randomly chosen and transmitted.

B. Non-real-time multivariate approach

This technique utilizes the cross-correlation among different variables of the multivariate data sensed by a smart meter to decrease the dimensionality of the data. Once the dimensionality is reduced, for each stream/dimension, temporal compression is exploited. Since the data is fluctuating in real-world conditions, the vital parameters, namely temporal sparsity and a minimum number of required dimensions are computed for each batch. The algorithm operates in two steps as explained below.

1) Principal component analysis: PCA is a technique to reduce the dimensionality of the data by transforming it from n-dimensions to p-dimensions. With this operation, the strongly related features in the input and a major portion of the variance of the whole multivariate data are preserved in p-dimensions. From the input data matrix, eigen vector eigen values are computed. In the transformed space, eigen vector serves as an orthogonal basis. The basis vectors are dependant on the input data, unlike other transformation techniques. The principal components obtained are arranged in decreasing order of variance and are uncorrelated.

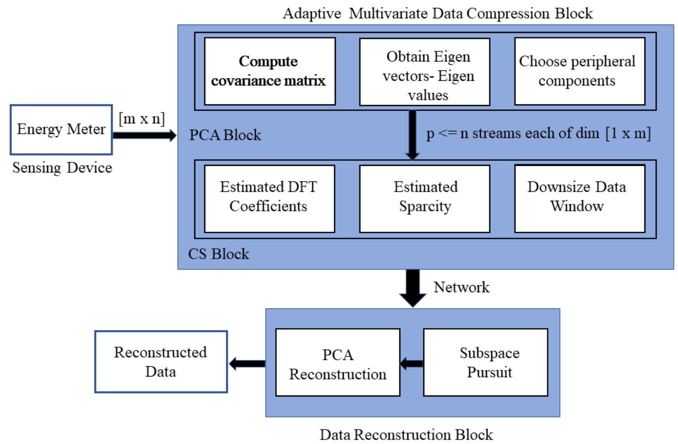


Fig. 2: Flow diagram of multivariate approach.

2) Compressive sensing: It is a technique that provides a compressed representation of data without much loss of information content. The signal is reconstructed by finding a solution to an under-determined linear system. The condensed representation provides saving in data storage and transmission. The properties of sparsity and incoherence are used in mathematical algorithms to reconstruct the signal from a few number of measurements. For reconstruction purpose, Subspace Pursuit Algorithm [13] has been used for data reconstruction since it is accurate and has low computational complexity.

3) Adaptive multivariate data compression (AMDC): Adaptive multivariate data compression [6] executes in two stages at the transmitter, i.e., the source IoT node. It processes data in batches. It operates on 2-D data, where one dimension represents the parameter type, and the other dimension represents the time scale. PCA is applied to decorrelate the input variables. Using the eigen value eigen vector combination, the principal components are obtained. Since this operation preserves the variance in the data in a few dimensions, so only those principal components are considered for the purpose of recovering the data depending on a predefined threshold. Then each stream of principal component is sent for temporal compression where sparsity is determined by the number of DFT coefficients that contribute at least 99.99% energy of the samples. The adaptive part of the algorithm in each stage is the number of principal components, and the number of transmitted samples are determined for each batch. Reconstruction at the receiver is also a two-stage process. The compressed version of the principal components is recovered using Subspace pursuit algorithm to yield the principal components. These are further fed to a PCA reconstruction block, which yields the actual data with reasonable accuracy.

IV. Prototype Implementation of Data-Smart IoT Module in Energy Meter

This section covers the hardware components that were selected for building the prototypes. Cost-effectiveness was the primary factor in the selection of a particular hardware module. Fig 4 shows the experimental setup of different prototypes.

A. Prototype I: Implementation on R-Pi

At first we have started implementing the smart meter data pruning algorithm in a R-Pi based platform with WiFi connectivity. The primary components are described below. Table I portrays the cost estimation for this implementation. Additional cost of SIMCOM NB-IoT HAT for NB-IoT connectivity of R-Pi is about USD 25.0, which is not included in Table I.

- Enersol MFR 2810 with RS485 port: This power meter is cost-effective, easy to operate, and compact in size. It is capable of measuring basic parameters that are required to monitor an electrical installation. The collected parameters with the meter are voltage, current, frequency and apparent power. This meter provides a RS 485 Modbus RTU compliant port whose data from various registers can be read. A programming manual [13] is provided to configure the various parameters of the meter, including the Modbus slave configuration.
- R-Pi development board: The R-Pi is a cheap and small-sized single-board computer that runs a full-fledged OS. It provides connectivity to common peripheral devices like a monitor, a standard mouse and keyboard and runs Raspbian OS, a variant of Linux OS. It also provides GPIO pins to control electronic devices for physical computing and develop IoT devices. It also provides UART pins for serial communication and having inbuilt WiFi.

TABLE I: Unit cost of prototypes

Unit cost of prototype I	
Component	Price (USD)
Enersol MFR 2810 3 phase meter with RS 485 port	48.86
R-Pi Development Board	38.96
PVC Meter Enclosure	6.72
RS 485 - USB Converter	2.55
Total	97.10
Unit cost of prototype II	
Enersol MFR 2810 3 phase meter with RS 485 port	48.86
STM32 Nucleo-64 Board	14.79
PVC meter enclosure	6.72
RS 485 - RS 232 Bi-dir converter	4.44
SIMCOM NB-IoT HAT	25.54
Total	100.35
Unit cost of prototype III	
Enersol MFR 2810 3 phase meter with RS 485 port	48.86
Smart IoT Node	26.89
PVC meter enclosure	6.72
RS 485 - RS 232 Bi-dir Converter	4.44
Total	86.91

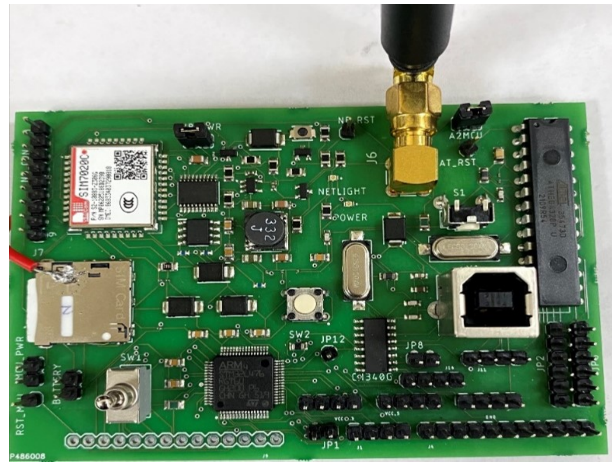


Fig. 3: Designed data-smart IoT module.

Design decision: Various low power wide area network (LPWAN) technologies like NB-IoT, SigFox, LoRa WAN are gaining the importance because of their extreme low power consuming nature during communication. As a part of IoT, implementation of these technologies are applied in smart meters also [14]. NB-IoT based connectivity will improve the energy efficiency of smart meter which will serve both the goals of cost effectiveness and commercial sustainability [15]. As our designed module will be in sleep mode for maximum time period, we will prefer to use Narrow Band-IoT (NB-IoT) communication protocol over WiFi for its low power budget during sleep and idle mode.

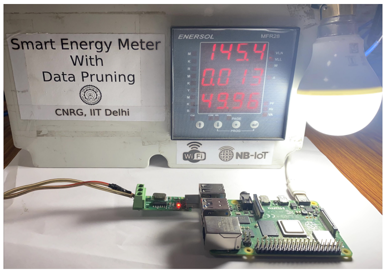
B. Prototype II: Implementation on microcontroller development board

Due to the high power consumption of R-Pi, we have revised our implementation methodology and come up with a more energy efficient, commercially acceptable design incorporating STM32 microcontroller exploiting the NB-IoT technology as communication medium. Cost estimate of this prototype incorporating low power NB-IoT is shown in Table I.

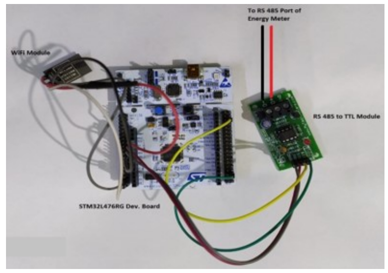
- NUCLEO - L476RG development board - The STM32L476RG [16] is a microcontroller based on the high-performance Arm Cortex 32-bit RISC core. It is an ultra low power microcontroller operating at a frequency of up to 80 MHz.
- SIMCOM 7020E NB-IoT module - The SIM7020 is a Multi-Band NB-IoT module. This module provides a rich interface like UART, GPIO etc. for issuing AT commands. It can be easily integrated with the user hardware modules. Therefore, it reduces the investment of the customer and also shortens the time-to-market.

C. Prototype III: Our designed data-smart IoT module

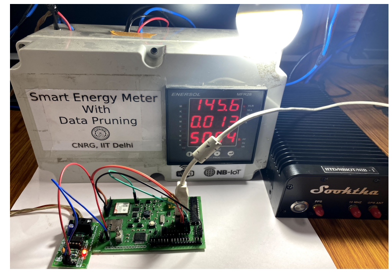
To reduce the energy consumption even further and provide a complete commercial cost effective solution, we have designed the low energy consuming smart IoT node. The node comprises of mainly two modules: low power STM32L476RG



(a) Setup with R-Pi

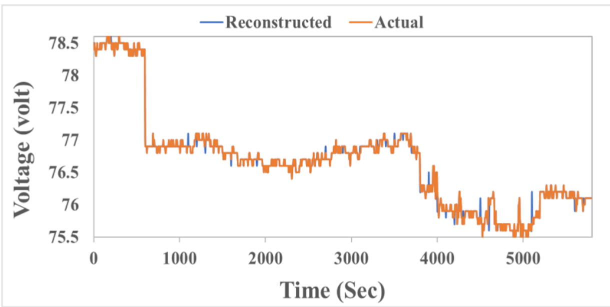


(b) Setup with STM32 Neucleo board

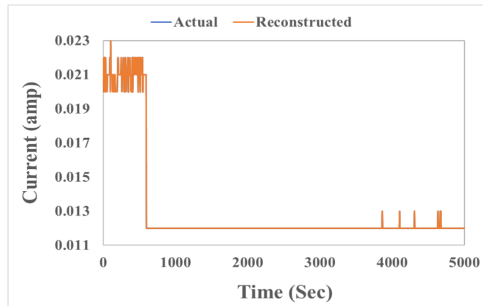


(c) Setup with designed smart IoT node

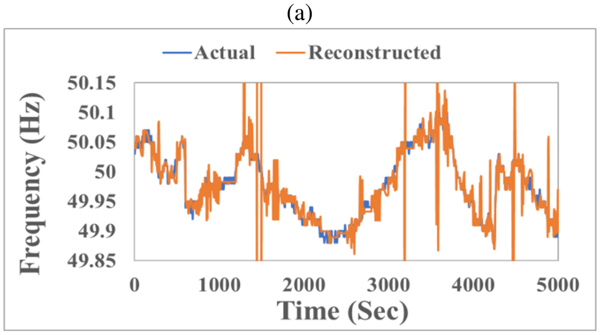
Fig. 4: Experimental setup of different prototypes.



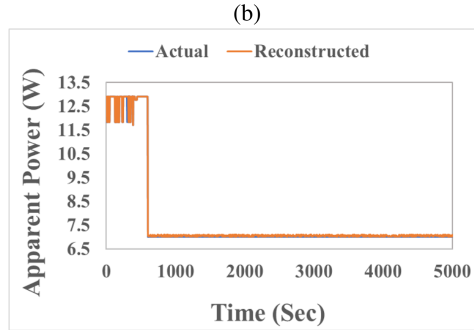
(a)



(b)



(c)



(d)

Fig. 5: Actual and reconstructed data collected from the cloud having overall RMSE 0.0004.

microcontroller to program the device and SIM7020C NB-IoT communication module. Once connected with the meter, the node will process the data collected from the meter and send it to the cloud platform through NB-IoT connectivity. Figure 3 shows the node. Table I estimates the cost of the smart meter with developed pruning subsystem.

V. Experimental Results and Discussion

A. Reconstruction performance

For the proper reconstruction of data, RMSE value should not cross the permissible threshold. The error threshold different kind of application of the smart meter data. RMSE below 0.2 is considered to be acceptable as described in [17]. In the present work, the RMSE value is 0.0004 for reconstruction of all variables. Fig 5 shows the comparison between actual and reconstructed (after decompression) data at the receiver. In can be observed that, in all cases, the actual and reconstructed data are overlapping which indicates a high reconstruction accuracy.

B. Energy efficiency

The designed smart IoT node exploits in many fold energy reduction of the smart meter implementation. By incorporating the two data pruning algorithms in the existing metering system we are able to reduce the processing energy significantly. We can observe from Table II that, implementing univariate and multivariate pruning algorithms on the designed smart IoT node, saves 19.7% and 36% more energy respectively than that of without any pruning algorithm. Further, the developed node is 99.09% energy sustainable than the R-Pi based counterpart.

C. Bandwidth saving

We have captured the plotting engine and showing it in Fig 6. We can observe that along with accurate reconstruction, we have achieved 31% bandwidth saving applying univariate data pruning approach [ref to III-A], however, it can go up

TABLE II: Energy consumption comparison

Univariate data pruning			
Hourly energy consumption (J)	Prototype I (R-Pi)	Prototype II (STM32)	Prototype III (Developed data-smart IoT module)
Without data intelligence	6832.92	63.45	62.15
With data intelligence	6769.77	61.96	49.89
Multivariate data pruning			
Without data intelligence	6514	59.13	58.79
With data intelligence	6448	54.79	37.54

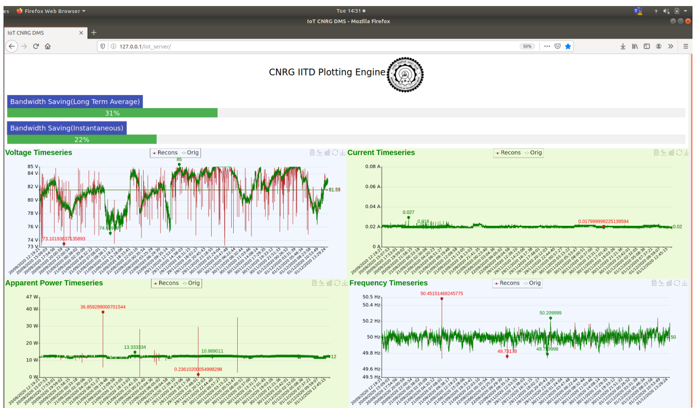


Fig. 6: Plotting engine showing actual and reconstructed data plots with bandwidth saving.

to 36% depending on the data variation profile. This we can further extend and can reduce the bandwidth requirement for multivariate data transmission by up to 98.5% using AMDC technique (refer to Section III-B3).

VI. Concluding Remarks

IoT has emerged as a leading-edge technology for automation, process monitoring, and system control. Hence the growth of IoT is undeniable. In this work, through the prototype implementation and testing studies we have highlighted that incorporating intelligence in IoT devices would make these devices smart in the context of data handling, thereby saving energy consumption as well as bandwidth footprint. Another contribution of this work is the design and implementation of a cost-effective data-smart IoT module in the example case of smart energy metering application. The benefit of the proposed data-smartness at the IoT nodes and the design principles of the smart data handling module is generically applicable in the other IoT applications dealing with non-real-time data.

Acknowledgement

This work has been supported by the Department of Telecommunication, Government of India, under the Grant No. 4-23/5G test bed/2017-NT, for building end to end 5G test bed, and by the Science and Engineering Research Board, Department of Science and Technology, Government of India, under the Grant No. CRG/2019/002293.

References

[1] T. N. Le, W.-L. Chin, D. K. Truong, T. H. Nguyen, and M. Eissa, "Advanced metering infrastructure based on smart meters in smart grid," *Smart Metering Technology and Services-Inspirations for Energy Utilities*, 2016.

[2] F. Clarizia, D. Gallo, C. Landi, M. Luiso, and R. Rinaldi, "Smart meter systems for smart grid management," in *Proceedings of IEEE International Instrumentation and Measurement Technology Conference Proceedings*. IEEE, 2016, pp. 1–6.

[3] G. Gaur, N. Mehta, R. Khanna, and S. Kaur, "Demand side management in a smart grid environment," in *Proceedings of IEEE International Conference on Smart Grid and Smart Cities (ICSGSC)*, 2017, pp. 227–231.

[4] S. Sreedevi, P. Prasannan, K. Jiju, and I. I. Lekshmi, "Development of indigenous smart energy meter adhering Indian standards for smart grid," in *Proceedings of IEEE International Conference on Power Electronics, Smart Grid and Renewable Energy (PESGRE2020)*. IEEE, 2020, pp. 1–5.

[5] S. Tripathi and S. De, "An efficient data characterization and reduction scheme for smart metering infrastructure," *IEEE Transactions on Industrial Informatics*, vol. 14, no. 10, pp. 4300–4308, 2018.

[6] M. R. Chowdhury, S. Tripathi, and S. De, "Adaptive multivariate data compression in smart metering Internet of Things," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 2, pp. 1287–1297, 2021.

[7] W. A. Khan, P. Das, S. Ghosh, M. R. Chowdhury, S. Tripathi, S. Kaur, S. Chatterjee, and S. De, "Smart iot communication: Circuits and systems," in *Proceedings of International Conference on COMmunication Systems NETworkS (COMSNETS)*, 2020, pp. 699–701.

[8] V. Gupta, S. Tripathi, and S. De, "Green sensing and communication: A step towards sustainable IoT systems," *Journal of the Indian Institute of Science*, vol. 100, no. 2, pp. 383–398, 2020.

[9] S. Tripathi and S. De, "Channel-adaptive transmission protocols for smart grid IoT communication," *IEEE Internet of Things Journal*, vol. 7, no. 8, pp. 7823–7835, 2020.

[10] T. Tony, P. Sivraj, and K. Sasi, "Net energy meter with appliance control and bi-directional communication capability," in *Proceedings of International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, 2016, pp. 2650–2653.

[11] W. Hlaing, S. Thepphaeng, V. Nontaboot, N. Tangsunantham, T. Sangsuwan, and C. Pira, "Implementation of WiFi-based single phase smart meter for Internet of Things (IoT)," in *Proceedings of International Electrical Engineering Congress (iEECON)*. IEEE, 2017, pp. 1–4.

[12] R. S. Sinha, Y. Wei, and S.-H. Hwang, "A survey on LPWA technology: LoRa and NB-IoT," *ICT Express*, vol. 3, no. 1, pp. 14–21, 2017.

[13] W. Dai and O. Milenkovic, "Subspace pursuit for compressive sensing signal reconstruction," *IEEE transactions on Information Theory*, vol. 55, no. 5, pp. 2230–2249, 2009.

[14] G. Wibisono, S. G. Permata, A. Awaludin, and P. Suhasfan, "Development of advanced metering infrastructure based on LoRa WAN in PLN Bali toward Bali Eco smart grid," in *Proceedings of Saudi Arabia Smart Grid (SASG)*. IEEE, 2017, pp. 1–4.

[15] P. Das, S. Ghosh, S. Chatterjee, and S. De, "Energy harvesting-enabled 5G advanced air pollution monitoring device," in *Proceedings of IEEE 3rd 5G World Forum (5GWF)*. IEEE, 2020, pp. 218–223.

[16] S. Website. Stm32l476rg. [Online]. Available: https://www.st.com/en/microcontrollers-microprocessors/stm32l476rg.html,

[17] J. Qin, Q. Zhao, H. Yin, Y. Jin, and C. Liu, "Numerical simulation and experiment on optical packet header recognition utilizing reservoir computing based on optoelectronic feedback," *IEEE Photonics Journal*, vol. 9, no. 1, pp. 1–11, 2017.