Digital Twin
of
Manufacturing Systems

Sujit Rokka Chhetri
Sina Faezi
Mohammad Abdullah Al Faruque

Center for Embedded and Cyber-Physical Systems
University of California, Irvine
Irvine, CA 92697-2620, USA

{schhetri, sfaez, alfaruqu}@uci.edu

CECS Technical Report CECS TR#17-07
November 27, 2017
LIST OF FIGURES
1 Digital Twin of an Additive Manufacturing System. ................................. 5
2 Constant lighting environment for measuring the quality of surface texture ($K_1$). ................................. 6
3 Last three steps (left to right) for measuring the quality of surface texture $K_1$. ................................. 6
4 Dynamic Data-Driven Application Systems Enabled Digital Twin of Additive Manufacturing Cyber-Physical System. ................................. 7
5 Classification Scores for various Sensor Position ........................................ 8
6 Experimental Setup for the Digital Twin. .................................................. 9
7 Data acquisition setup for the experimental analysis of Digital Twin models. ................................. 9
8 A sample snapshot of the data collected from the sensors. ................................. 10
9 Test 3D object for Digital Twin Experiment. ............................................... 11
10 Segmentation of Test 3D object for Digital Twin Experiment. ................................. 12
11 Printed Test Objects. ........................................................................... 13
12 Various test objects printed with varying flow-rates. ................................. 13
13 Aliveness test result for Digital Twin model predicting thickness KPI for the floor and the wall. ................................. 14
14 Feature Re-ranking results for Digital Twin predicting Floor and Wall Thickness ................................. 15
15 Aliveness test result for Digital Twin model predicting Surface KPI for the top and bottom surface of floor and front and back surface of wall. ................................. 15
16 Surface textures (for segment number 2) with flow rate variation and the case of unaccounted trench. ................................. 16
17 Feature Re-ranking results for Digital Twin predicting Texture for Top and Bottom Surface of the Floor ................................. 16
18 Feature Re-ranking results for Digital Twin predicting Texture for Front and Back Surface of the Wall ................................. 17

LIST OF TABLES
I List of sensors used for monitoring the 3D printer and its surrounding environment. ................................. 9
II Summary of Environmental and Aging Degradation Parameters ................................. 11
Abstract

The digitization of manufacturing systems is at the crux of the next industrial revolutions. The digital representation of the “Physical Twin,” also known as the “Digital Twin,” will help in maintaining the process quality effectively by allowing easy visualization and incorporation of cognitive capability in the system. In this technical report, we tackle two issues regarding the Digital Twin: (1) modeling the Digital Twin by extracting information from the side-channel emissions, and (2) making sure that the Digital Twin is up-to-date (or “alive”). We will first analyze various analog emissions to figure out if they behave as side-channels, informing about the various states of both cyber and physical domains. Then, we will present a dynamic data-driven application system enabled Digital Twin, which is able to check if it is the most up-to-date version of the Physical Twin.

Index Terms

Digital Twin, Cyber-Physical Systems, Digitization, Additive Manufacturing, Machine Learning, Sensor Fusion, Dynamic Data-Driven Application Systems

I. INTRODUCTION

Cyber-physical manufacturing systems have various cyber processes interacting with the physical domain components through a communication network [1]. The processes in the cyber domain can be monitored for each clock cycle of the computing and the communicating component. This allows us to accurately predict the behavior of the cyber domain components. However, in the physical domain the same is not true. Monitoring physical domain components, and predicting the physical system behavior are faced with a myriad of challenges [2], [3]. One of the major challenges arises due to the fact that the physical domain continuously interacts with environment, humans, etc., which changes the states of the physical components [1], [4]–[9]. Predicting these interaction and the effects of environment has been one the major research topics in the manufacturing industry [10]–[13]. This is because these may affect the quality of the products being produced.

To overcome the problem of process control, among many others, digitization of the physical domain can be one of the solutions [2], [14]–[16]. This real-time digital representation of the physical domain, also known as ”The Digital Twin” [3], [17], [18], will make it is easier for manufacturers to accurately predict the future system performance and plan accordingly as well [2], [19], [20]. The future trend will be in using the capability of Internet of Things in producing massive amount of data to create a Digital Twin that interacts with the cyber domain [21], [22]. In fact, Gartner predicts the digitization will be the top trend in coming years, and the Digital Twin has been placed in the innovation trigger phase of emerging technologies in the Hype Cycle [23]. The concept of Digital Twin was first used by NASA, whereby they wanted a Digital Twin replica of the physical system used for the space exploration [24]. Since then, various works have been done to create the Digital Twin of the Physical Twin [17], [25]. Authors in [26] have created Digital Twin, an ultrahigh fidelity model of individual aircraft, to predict the structural integrity and the life of the aircraft structure. Company like IBM are providing sensor systems for creating the Digital Twin of buildings [27]. In fact, technology solutions are hitting the market for enabling creation of Digital Twin. Ansys [28], released their technology solution for building the twin, with an example for creating the Digital Twin for pumps. Honeywell has recently presented a connected plant concept with a solution to bring all process domain knowledge to a digital twin in the cloud [29]. Digital Twin technology is also being used for lifecycle engineering asset management in [30], [31]. Authors in [32], [33] provide analysis for research needs and current status for the building blocks of the first generation of the Digital Twin for additive manufacturing (or a 3D printer) system. In summary, the benefits of building and updating a Digital Twin of a physical system has recently been acknowledge both by researchers and industries. However, the Digital Twin technology is still in its infancy, and it faces various research challenges such as:

- Understanding what variables in the physical domain can be extracted.
- Knowing how to select the number and position of sensors either during design time or during the usage time (for legacy systems).
- Understanding how a Digital Twin model can be developed given the constraints of resources (such as sensors).
- Figuring out when to update the Digital Twins (as lightweight as possible to meet the resource constraints), to make sure that they can accurately predict the system performance.

In order to solve these research challenges, in this research project, we provide following various solutions:

- Present analysis of physical domain signals that can be extracted from cyber-physical manufacturing systems (Section V-A.1).
- Perform feature engineering on signals, and data driven modeling for creating the Digital Twin (Section IV-B).
- Use dynamic data-driven application systems for keeping the digital twin up-to-date and lightweight (by re-ranking and selecting the most prominent features for building the Digital Twin) (Section IV).

The first major contribution is the analysis of the analog emissions to extract information from the physical domain. These analog emissions may behave as side-channels by leaking information [34]. Side-channels have been used to extract information about the cyber domain (such as secrets keys), without using brute force or weakness of the cryptographic algorithm, but rather using the physical implementation system [35]. Researchers have recently demonstrated that side-channel analysis can
be performed at a system level to extract valuable information about the cyber domain [36]–[43]. Since the cyber domain information is manifested in the physical domain, these side-channels also carry information about the cyber domain. Hence, this information can be used to build an effective Digital Twin of the system. The second contribution is providing a methodology for keeping the Digital Twin alive. This methodology is based on Dynamic Data-Driven Application Systems (DDDAS) [44] concepts. The DDDAS concept is used in influencing the data-driven models (in our case the Digital Twins), by providing a dynamic feedback in updating the models based on the real-time data from the physical domain.

The goal of this project is to demonstrate a cognitive capability of manufacturing systems using the Digital Twin. The key technical challenge of modeling and updating a living Digital Twin of a Physical Twin is solved with the use of dynamically steered sensor data processing (selection and fusion), which uses concepts of DDDAS. This project aims to provide scientific contribution in closing the loop between the Physical Twin and its Digital Twin for maintaining the most up-to-date virtual representation of the cyber-physical manufacturing system.

II. DIGITAL TWIN OF ADDITIVE MANUFACTURING CPS

![Diagram of Digital Twin](image)

**Environmental Effects**

+ Aging \((B_1, B_2, ..., B_n)\)

**Process Parameters** \((\alpha)\)

**Design Parameters** \((\beta)\)

**Digital Twin**

**Key Performance Indicators (KPIs)**

- Surface \((K_1)\)
- Dimension \((K_2)\)

**Analog Emissions**

- Acoustic \((A_1)\)
- Vibration \((A_2)\)
- Power \((A_3)\)
- Magnetic \((A_4)\)

Fig. 1: Digital Twin of an Additive Manufacturing System.

A. Digital Twin Model

The major contribution of this project is the creation of the Digital Twin based on the dynamic feature selection using side-channel emissions. Before building the Digital Twin, we need to setup an expected outcome or a use case for the Digital Twin. An optimistic version of the Digital Twin would be able to demonstrate and predict every possible physical state of the system. However, to narrow down the scope of this projects, we set two objectives for the Digital Twin:

- Based on the status of the physical components, predict the KPIs such as surface texture and dimension of the 3D object that will be printed in the future.
- Be able to know when the digital representation of the physical components are not up-to-date for the KPI predictions.

With these narrowed down scopes, let us explain how the Digital Twin can be modeled in an additive manufacturing cyber-physical systems. In Figure 1, we present the Digital Twin representation. It takes into consideration the various environmental interactions and aging phenomena into consideration (represented by random variables \(B_1, B_2, ..., B_n\)), the process (random variable \(\alpha\)) and design parameters (random variable \(\beta\)), the historical and current analog emissions from the systems (represented by random variable: Acoustics \((A_1)\), Vibration \((A_2)\), Power \((A_3)\), Magnetic \((A_4)\), etc.), and predict the KPI, such as surface texture (represented by a random variable \(K_1\)) and dimension (represented by a random variable \(K_2\)). In doing so, it explains the effect of environmental parameters \((B_1, B_2, ..., B_n)\) on process and design parameters \((\alpha, \beta)\) which in return affects the KPIs. With this backdrop, the Digital Twin can be modeled using a machine learning algorithm that explains the relation between analog emissions, process parameters, environmental factors, and the KPIs.

B. Key Performance Indicators (KPIs)

The two KPIs, used in measuring the performance of the additive manufacturing systems are as follows:

1) Surface Texture \((K_1)\): The quality of surface texture of the printed object is one of the most sensitive performance indicators that is affected by different design and process parameters. For instance, slightly higher temperature of the base plate or extruder can cause distortion on the surface by affecting the property of the filament being deposited. Moreover, the slice thickness of the 3D object and the road width of the segment being printed on the \(XY\)-plane also cause the surface texture to vary in quality. For quantifying the surface quality, we explored an innovative metric called dispersion of directionality \((K_1)\). It is a heuristic metric calculated by performing various image processing algorithms on the surface image. **We would like to emphasize that**
other standard metric can be used with our methodology (as the methodology is independent of the metric used to measure the KPI), and in fact may improve the accuracy of the Digital Twin. The steps to calculate $K_1$ are as follows:

1) **Create a constant lighting environment:** As shown in Figure 2, we use an empty box with a small opening for fitting a DSLR camera on the top (this camera can be replaced with a low cost camera as well). This enclosure prevents the time varying external light from entering and affecting the surface texture measurement. To maintain a constant lighting, we used two similar light sources, around 10 lumen of luminous flux, placed directly opposite to each other to provide homogeneous lighting on the surface of the 3D object. Moreover, guide lines are drawn to always place the 3D object on the same location for the surface texture measurement.

2) **Remove the background:** Since the 3D objects have a distinct color (green) compared to the background (brown), we first transform the image taken by the camera from the RGB to the Lab color space [45]. Lab color space consists of three dimensions: $L$ for lightness of the image, and $a$ and $b$ for the color opponents green-red and blue-yellow, respectively. After this transform, we choose the value of $a$ which eliminates the shadows and the brown colored background. Then, we perform a constant threshold to create a mask, which matches the green colored object. This mask is then applied to the image to eliminate the background.

3) **Surface Division:** The image is divided into either 4 by 4 or 16 by 16 equal parts for aiding the process of mapping the surface texture $K_1$ value to its corresponding analog emissions.

4) **2D Discrete Fourier Transform:** The discrete transform ($F[k,l]$) is calculated for the image of the 3D object’s surface ($f[m,n]$) (see Figure 3 (b)), using Equation 1, where, $M$ and $N$ are the height and the width of the image calculated in the step 3. The maximum values for $k$ and $l$ are $M$ and $N$ as well.

$$F[k,l] = \frac{1}{MN} \sum_{m=0}^{M-1} f[m,n]e^{-j2\pi(\frac{k}{M}m+\frac{l}{N}n)}$$ (1)

5) **Directionality histogram calculation:** Based on the value of $F[k,l]$, calculated in the previous step, a directionality histogram is calculated using the approach mentioned in [46] for faster calculations (see second image in Figure 3).

6) **Fit Normal Distribution:** A normal curve is then fitted to the histogram (see Figure 3 (c)) obtained in the previous step, and the corresponding distribution parameters are calculated. Out of these parameters, standard deviation $\sigma$ is the dispersion metric we use to measure the directionality and hence the surface texture $K_1$.

![Fig. 3: Last three steps (left to right) for measuring the quality of surface texture ($K_1$).](image)

The heuristic dispersion metric used for surface texture measurement $K_1$ does not have a linear relation with number of the lines in the image or the amount of disorientation in the image. In order to prove the effectiveness of dispersion metric for surface texture measurement, we asked two people to sort 28 surface images of the 3D objects printed by the fused deposition
modeling based additive manufacturing system. These surface textures were varied randomly with various levels of surface quality. This sorting was then compared to the sorting order created using the values of dispersion of directionality calculated using our algorithm. The result was a similar sorting order. However, due to the subjective nature of quality perception, further work is needed to compare the result of the standard metric and the dispersion of directionality for surface texture measurement. Nonetheless, in this work, dispersion of directionality is used as a KPI for surface texture measurement, due to its ability to represent the varying road thickness of the filament deposited on the XY-plane, presence of non-directional surface pattern, etc. Moreover, the use of standard metric will only make the Digital Twin modeled using our approach more accurate.

2) Dimension ($K_2$): The dimensional accuracy is affected by various process and design parameters. The process parameters that affect $K_2$ are build environment temperature, filament feed-rate, nozzle temperature, etc. Whereas, the design parameters that affect the $K_2$ are road width, slice thickness, air gap, build orientation, raster patterns, etc. Various environmental factors such as room temperature, humidity, vibrations, etc. can influence these process and design parameters, which in return affect $K_2$. Since the resolution of the 3D printer used in our experiment is about 12.5 micron, 12.5 micron, and 2.5 micron in X, Y, and Z directions, respectively, micrometer is used to measure the dimension of the 3D objects.

### III. KEEPING DIGITAL TWIN UPDATED

![Fig. 4: Dynamic Data-Driven Application Systems Enabled Digital Twin of Additive Manufacturing Cyber-Physical System.](image)

The Digital Twin model explains the relationship between KPIs, analog emissions, environmental factors, process parameters and design parameters. In order to keep the Digital Twin updated, these models must have to be continuously updated based on the run time difference ($\delta$) between the predicted KPIs and the measured KPIs (as shown in Figure 4). This $\delta$ will give the Digital Twin an intuition of liveliness of the Digital Twin and enable it to have a cognitive ability. If the error between the predicted and the real KPI value is large then the Digital Twin will be assumed to be outdated. This in return will trigger a feedback to the dynamic feature selection algorithm to re-rank and fuse the feature values to measure the relation between the most recent texture value and the analog emission values calculated, while only selecting the features that are necessary (thus keeping the Digital Twin model lightweight). The steps for updating the Digital Twin (as shown in Figure 4) are as follows:

1) The product Digital Twin (3D object) is given to the Digital Twin of the 3D printer.
2) Digital Twin of the 3D printer predicts the KPIs ($K_1$ and $K_2$).
3) The 3D Object is sent to the physical twin for production only when the predicted KPIs are within tolerable range.
4) Actual KPIs are measured from the printed 3D Object if it is printed.
5) The cognitive algorithm measures the difference between the actual and predicted KPIs.
6) If the $\delta$ is large, feedback is sent to the Digital Twin and the feature processing algorithms.
7) Dynamic feature processing algorithm re-ranks and fuses the features to get the most up-to-date Digital Twin model.

The feedback to data-driven Digital Twin model is activated by the analog emission data gathered from the sensors. This data represents the current physical state of the system (for example, mechanical degradation may result in variation of the analog emission pattern).

### IV. BUILDING DIGITAL TWIN

#### A. Sensor/ Emission Modality Selection

3D printer consist of set of actuators, mechanical moving parts, heating elements, and a controller board. An actuator under the load consumes energy, vibrates, emits electromagnetic waves, and produces audible sound. The heating elements of the
3D printer, which are embedded inside the nozzle(s) and the base plate, consume electrical energy and convert it into thermal energy. Based on these facts, we decided to monitor the physical domain by acquiring acoustic, electromagnetic, vibration, power, humidity, and temperature data, and analyzed them for amount of information revealed about the cyber domain and physical domain states of the 3D printer. The validation for the analog emission modality selection is experimentally validated through the accuracy of the Digital Twin models.

B. Feature Engineering

After selecting the emissions to be monitored in the physical domain, various features have to be extracted to reduce the size of the raw data collected, and to improve the performance of the machine learning models (or the data-driven models) used to create the Digital Twin. In this section we will briefly describe the various features extracted in various domains.

1) Time Domain: The various features extracted in time domain are Energy, Energy Entropy, Mean Amplitude, Maximum Amplitude, Minimum Amplitude, Median Amplitude, Mode of Amplitude, Peak to Peak features (highest peaks, peak widths, peak prominence, etc.), Root Mean Square values, Skewness, Standard Deviation, Zero Crossing Rate, Kurtosis, etc (Total 114). These are extracted for each emission. Each of these features capture various properties of the signal, explaining each one of them is out of scope of this report, and will be left for the future work.

2) Frequency Domain: Various analog emissions are monitored in the physical domain. Each of them have different frequency range and characteristics. To capture all of the characteristics for all the signal, various frequency domain signals are analyzed.

- Frequency Characteristics and Short Term Fourier Transform (STFT): The frequency characteristics analyzed are based on the short term windows and various characteristics of the frequency domain such as: Mean Frequency, Median Frequency, Signal to Noise Ratio, Power Bandwidth, Spectral Centroid, Spectral Entropy, Spectral Flux, Spectral Roll Off, etc (82 in Total).

- Frequency Characteristics and Continuous Wavelet Transform (CWT): The challenge with short term Fourier Transform based features is that the trade off between time and frequency. The window frame (in time domain) and the resolution of frequency domain features are highly dependent, and one has to be compromised for the other. Instead of compromising these we have also analyzed continuous wavelet transform in frequency domain, and analyzed various characteristics of the transform (in Total 58). Based on the result of the Digital Twin prediction and feature ranking, in future, the continuous wavelet transform will be used to calculate the discrete wavelet transform with specific approximation and detailed coefficients.

C. Sensor Positioning

In order for us to determine the position of the sensors selected to acquire the data from the physical domain, we performed profiling of the analog emissions from each modality. For this, a classifier is modeled to classify various cyber domain data (G/M-codes), such as movement in each axis. Then, various location around the 3D printer (9 locations for each modality except power, humidity and temperature) were selected for measuring the analog emission, and the corresponding classification scores. A Matlab based graphical user interface is created to analyze the classification scores and the corresponding feature ranking scores for analyzing and finalizing the sensor positions. In Figure 5, the result of the sensor position GUI and the four channel’s classification scores are presented. It can be seen from the figure that for channel 1, position 9 gives the highest classification accuracy, whereas for channel 2, channel 3 and channel 4, the position most favorable for higher accuracy are 6, 5, and 2, respectively.

![Sensor Position Data GUI in Matlab](image1)

![Classification Scores](image2)

Fig. 5: Classification Scores for various Sensor Position
D. Data-Driven Models

For creating the data-driven models of the Digital Twin, we explored various machine learning algorithms such as Gradient Boosted Regressor [47], Decision Tree regressor [48], K nearest Neighbor Regressor [49], AdaBoost Regressor [50], etc. These models were used to model the relationship between the design and process parameters and the analog emissions, and the KPIs.

V. EXPERIMENTAL SETUP

The experimental setup for modeling and updating the Digital Twin is shown in Figure 6. As a test case, we have selected fused deposition modeling based additive manufacturing cyber-physical systems. The various components of the experiments are explained as follows:

TABLE I: List of sensors used for monitoring the 3D printer and its surrounding environment.

<table>
<thead>
<tr>
<th>Quantity</th>
<th>Modality</th>
<th>Sensor Model</th>
<th>Interface</th>
<th>Outputs</th>
<th>Sampling Rate (kHz)</th>
<th>Sensitivity</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>Vibration</td>
<td>ADXL326</td>
<td>Analog</td>
<td>3</td>
<td>1.6 (X,Y), 0.55 (Z)</td>
<td>57 mV/g</td>
</tr>
<tr>
<td>3</td>
<td>Magnetic</td>
<td>HMC5883L</td>
<td>Digital (I2C)</td>
<td>3</td>
<td>0.16</td>
<td>2 mG</td>
</tr>
<tr>
<td>1</td>
<td>Current</td>
<td>PICO TA018</td>
<td>Analog</td>
<td>1</td>
<td>20</td>
<td>100 mV/Amp</td>
</tr>
<tr>
<td>3</td>
<td>Acoustic</td>
<td>AT2021</td>
<td>Analog</td>
<td>1</td>
<td>20</td>
<td>11.2 mV/Pa</td>
</tr>
<tr>
<td>1</td>
<td>Acoustic/Vibration</td>
<td>AKG C411 III</td>
<td>Analog</td>
<td>1</td>
<td>18</td>
<td>2 mV/Pa</td>
</tr>
<tr>
<td>1</td>
<td>Temperature</td>
<td>LM35</td>
<td>Analog</td>
<td>1</td>
<td>1.5</td>
<td>10 mV/°C</td>
</tr>
<tr>
<td>1</td>
<td>Humidity</td>
<td>AM2001</td>
<td>Analog</td>
<td>1</td>
<td>0.0005</td>
<td>0.1%RH</td>
</tr>
</tbody>
</table>

A. The Test-bed

The test-bed consists of an Ultimaker 3 3D printer, a set of sensors with analog and digital interface, a Data Acquisition (DAQ) device, two Arduino Uno microcontroller boards coupled with MCP4725 for digital to analog conversion (DAC) purpose, and a personal computer for managing the acquired data. The details of the test-bed are as follows:

1) Sensors: In this project, as presented in Table I, we monitor the 3D printer and its surrounding environment using a total of 25 sensors. Maximum sampling rate, sensitivity, and cost are the important factors that determine type of sensor selected for each modality. We have used three accelerometers for measuring the vibration in the system. These accelerometers have a sampling rate of 1 kHz. A contact microphone is used to measure the acoustic noise and the high frequency vibration from the printer. Three microphones with sampling rate of 20 kHz are used to measure the acoustic emissions in the audible range.
from the 3D printer. ZOOM TAC-8 phantom power supply and amplifier is used to condition the acoustic signal before feeding it to the DAQ. The electromagnetic field intensity variation caused by the stepper motors of the 3D printer cannot be captured without using high precision EM sensors. Instead, we use three compass sensors, which are designed to sense the Earth’s magnetic field. By using these magnetic sensors, we measure the fluctuation in the magnetic field of field of Earth caused by the moving metallic parts of the 3D printer. The sampling rate of the compass is around $270 \text{ Hz}$. The humidity and the temperature of the room vary slowly over time. Moreover, the KPIs do not change drastically within the $\pm 5\%$ fluctuation of humidity, and $\pm 1^\circ \text{C}$ change in the temperature. Hence, any sensor satisfying these properties would be sufficient for monitoring the analog emissions from these modalities.

2) Data Acquisition: NI USB-6343 OEM is used for data acquisition (DAQ), see Figure 7. It has 32 analog inputs. For DAQ, with the increasing number analog inputs the overall sampling rate decreases. This limits the number of analog signals that can be monitored with high sampling rate. However, for the 25 analog inputs used in our experiment, the resolution of the DAQ is 16 bits for the data with 20,000 Samples/Second sampling rate. These resolutions are sufficient for the acoustic emissions, and surpass the requirements for other emissions. Since the DAQ takes analog signal, an arduino board coupled with MCP4725 is used to convert the sensor’s digital data to analog form. This is done to synchronize all the 25 channels, and maintain coherent sampling and data resolution. The Arduino board reads the data from the sensor using I2C interface and sends it to MCP4725 over the I2C for conversion to analog form. This conversion is necessary for the magnetic sensors. There are three magnetic sensors, each measuring magnetic fluctuation in X, Y, and Z direction. This results in a total of nine signals given by the three magnetic sensors. Since the MCP4725 boards share the same I2C addresses, two TCA9548A I2C Multiplexers are used to access them separately. According to our measurements, this setup can convert more than $6 \times 275$ digital samples to analog signal every second, which is more than enough for converting the $3 \times 170$ samples generated by the compass sensors.

3) Data Synchronization: The DAQ used in this project assures synchronization of all the sensors’ data with each other (see Figure 8 for snapshot of data collected from the DAQ). However, the analog signals collected from the sensors should be mapped with G-code for building the Digital Twin. For being able to segment the G-code, the 3D printer firmware is modified to send every G-code (along with the time-stamp with accuracy in the range of milliseconds) right before execution to the host’s IP address. The port used for the communication is 5000. A Python code running on the host side (the desktop computer) is made to continuously listen at port 5000, waiting for the printer to start printing. Once the host receives the first G-code, it...
first saves the 3D printer’s clock data (which allows us to synchronize the G-code with the DAQ data). It then starts saving the sensors data (in .tdms format) sent from the DAQ in chunks of \( \sim 55 \text{ MB} \).

![A diagram of a 3D object with dimensions labeled: X: 25 mm, Y: 25 mm, Z: 0.4 mm.](image)

**Fig. 9**: Test 3D object for Digital Twin Experiment.

### B. Test 3D Objects

For modeling and testing the aliveness of the Digital Twin, we have customized a benchmark model (see Figure 9). Using this benchmark model, we measure the surface texture \((K_1)\) and dimension variation \((K_2)\), map these KPIs with the corresponding analog emissions from the side-channels, and build the Digital Twin model. Our benchmark model has four surfaces: top and bottom for floor, and front and back for the wall. Hence, we measure the surface texture of all four surfaces. For dimension, we measure the width and the breadth of the floor, the width and the height of the wall, and the thickness of both the floor and the wall (see Figure 9).

**TABLE II**: Summary of Environmental and Aging Degradation Parameters

<table>
<thead>
<tr>
<th>Degradation Parameter</th>
<th>Corresponding ( \alpha ) and ( \beta )</th>
<th>Environmental&amp;Aging Effects</th>
<th>( K_1, K_2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( B_1 ) Nozzle Temperature (( \alpha ))</td>
<td>(1) Sensor Malfunction, (2) Extreme environmental temperature, (3) Heating element malfunction degradation</td>
<td>( K_1, K_2 )</td>
<td></td>
</tr>
<tr>
<td>( B_2 ) Filament Feedrate (( \alpha ))</td>
<td>(1) Slippage, (2) Worn out rollers due to mechanical degradation, (3) Variation of filament thickness</td>
<td>( K_1, K_2 )</td>
<td></td>
</tr>
<tr>
<td>( B_3 ) Acceleration of Stepper Motors (( \alpha ))</td>
<td>(1) Rust, (2) Vibration due to loose components, (3) Vibration due to mechanical degradation</td>
<td>( K_1, K_2 )</td>
<td></td>
</tr>
<tr>
<td>( B_4 ) Power Outage (( \alpha ))</td>
<td>(1) Short circuit of electronic components, (2) PCB failure, (3) Power supply failure</td>
<td>( K_1, K_2 )</td>
<td></td>
</tr>
<tr>
<td>( B_5 ) Printer Alignment (( \alpha ))</td>
<td>(1) Shockwave, (2) Earthquake</td>
<td>( K_1, K_2 )</td>
<td></td>
</tr>
<tr>
<td>( B_6 ) Humidity (( \alpha ))</td>
<td>(1) Faulty HVAC</td>
<td>( K_1, K_2 )</td>
<td></td>
</tr>
<tr>
<td>( B_7 ) Slicing Thickness (( \beta ))</td>
<td>(1) Faulty Z-motor, (2) Erosion of translation screw</td>
<td>( K_1, K_2 )</td>
<td></td>
</tr>
</tbody>
</table>

### C. Data Collection

The main objective of the Digital Twin is to be able to predict the KPIs based on the environmental effects and aging. Collecting analog emissions, within a short period of time, that encompass the interaction between the system and the environment is a challenging task. Moreover, collecting the variation in the analog emissions due to aging requires collecting data for a long period of time. In order to carry out the experiment within a short period of time, we have performed the following tasks:

- First of all, a summary of environmental and aging effects and their corresponding impacts on the design and process parameters is listed (presented in Table II).
- Analysis of how each design and process parameter gets affected is performed.
- Then design and process parameters are varied to reflect the impact of environmental and aging effects.
- The KPIs, corresponding to the design and process parameter variation are measured.

From Table II, we were able to analyze which design and process parameters get affected by which environmental factors and aging degradation. For the purpose of the experiment, where we want to validate the DDDAS enabled Digital Twin, we selected the degradation parameter \((B_2)\) which can be reflected in terms of varying filament flow rate. The flow rate of the
system is changed to reflect the effects of environment and aging. To do this, the flow rate is changed from 20% to 200% of the optimal flow rate of the 3D printer with the step-size of 10%. In 3D printer, the optimal range for the flow rate lies in the range of 80% to 120%. The flow rate is the process parameter which is calculated using the equation:

\[ W \times H = A = \frac{Q}{v_{\text{feed}}} \]  (2)

Where \( W \) is the width and \( H \) is the height of the line-segment being printed on the XY-plane, \( Q \) is the constant volumetric flow rate of the material. \( Q \) is estimated based on die swelling ration, pressure drop value and buckling pressure of the filament. \( v_{\text{feed}} \) is the feed velocity of the filament and is calculated as:

\[ v_{\text{feed}} = \omega_r \times R_r \]  (3)

Where, \( \omega_r \) is the angular velocity of the pinch rollers, and \( R_r \) is the radius of the pinch rollers. Based on these values, the pressure drop is calculated as follows:

\[ P_{\text{motor}} = \frac{1}{2} \Delta P \times Q \]  (4)

Where, \( P_{\text{motor}} \) is the pressure applied by the stepper motors, \( \Delta P \) is the pressure drop. Hence, pressure applied by the motor needs to be maintained for the constant volumetric flow rate. However, this pressure needs to be less than buckling pressure calculated as follows:

\[ P_{cr} = \frac{\pi^2 \times E \times d_f^2}{16 \times L_f^2} \]  (5)

Where \( E \) is the elastic modulus of the filament, \( d_f \) is the diameter of the filament, and \( L_f \) is the length of the filament from the roller to the entrance of the liquifier present in the nozzle. It is evident from these equations that maintaining a constant flow rate depends on various parameters, and environmental or aging factors affecting any of these parameters will change the flow rate, causing changes in the KPIs.

While creating a Digital Twin in a manufacturing plant, this analysis need not be performed, and data from various analog emissions that have the likelihood of behaving as side-channels can be collected. This data can then be mapped to the KPIs to be able to model a Digital Twin that predicts the KPIs based on the varying process parameters. The variation in the process parameters are due to the environmental effects and the degradation due to aging.

**D. Data Segmentation**

![G-code Line Segments](image)

**Floor Segmentation**

**Wall Segmentation**

![Mapping Segmentation](image)

**Extracting G-codes Based on Segments**

**Fig. 10:** Segmentation of Test 3D object for Digital Twin Experiment.

The foremost task in any data-driven modeling is acquiring the labeled data for training the machine learning algorithms. In our case, we want to be able to map the KPIs with the corresponding \( \alpha \) and \( \beta \), and the analog emissions. In order to achieve this, we segmented the floor and wall of the test objects into various segments, mapped it to the corresponding G-codes of the object and acquired the timing data necessary to segment the analog emissions. For experimental purposes both the wall and floor region are segmented into 4 by 4 matrix segments. After segmenting the data, the various segments are grouped together for the purpose of mapping the corresponding KPIs with the analog emissions. For instance the Floor and Wall segments
numbers 3, 7, 8, and 13 (see Figure 10), are grouped to measure the surface texture and thickness, as they lie on the center region and have homogeneous surface texture and dimension. The reason for searching homogeneous segments is made clear through Figure 12. Due to the custom synchronizing code running inside the 3D-printer firmware, various non-homogeneous changes are created in the printed 3D object. These changes are more prominent on the outline of the object. Hence for the experimental verification purpose, in the work, we have discarded the non-homogeneous outer segments.

For each of the test objects, the total number of channels from which the data collected is 25. And for each channel the total number of features extracted is $114 + 82 + 58 = 254$. In our training model all the channel's features are fused together into a single matrix. Hence, the dimension of training data is $X_{\text{samples} \times 25 \times 254} = X_{\text{samples} \times 6350}$, where samples represents the total number of observations.

VI. SIMULATION AND RESULTS FOR DIGITAL TWIN MODELS

A. Digital Twin Models

With the objective of predicting the dimension and surface texture quality, we have created the following Digital Twin models.

- Two models, predicting the thickness of the wall and the floor.
- Four models predicting the surface texture of floor top, floor bottom, wall front, and wall back, respectively.

All of these models are created using gradient boosting based regressors [47], [51]. These are an ensemble of decision tree based regression models. By generating a new tree against the negative gradient of the loss function, this algorithms combines weak learners to control over-fitting. It is chosen due to its robustness against outliers and better predictive power against other regression algorithms. In this section, we will present some accuracy measures of the regression algorithm while training it for optimal values of flow rate ($80\%$ to $120\%$), and slowly incorporating degradation along the positive direction, i.e. ($130\%, 140\%, \ldots, 200\%$).

B. Aliveness

In order to check the aliveness of the Digital Twin, we devised the experiment as follows:

1) Train the Digital Twin with the optimal values of the flow rates ($80\%$ to $120\%$).
2) Assume an environmental degradation has caused the flow rate to vary to 200%. Predict the KPIs, with the Digital Twin trained with optimal flow rates.

3) Gradually incorporate the flow rates degradation from 130%, 140%, etc., all the way to 190%. This incorporation demonstrates a gradual update of the Digital Twin along the run-time performance of the 3D printer.

Through this experiment we will be able to measure two things: (1) the predicting capability of the Digital Twin modeled from side-channel emissions, and (2) cognitive capability enabled for DDDAS by checking the aliveness of the Digital Twin. The prediction capability is demonstrated in terms of accuracy of each of the models. By gradually adding the analog emissions in modeling the Digital Twin, we assume that we are adding the emissions that represent the physical status of the system that is closer to the current state. By doing so, we expect to see improvement in the prediction capability of the Digital Twin and lower δ value. The lower value of δ will signify that it is possible to monitor the δ and infer about the aliveness of the Digital Twin model.

Since we have used the gradient boosted trees, the feature re-ranking is performed based on the relative importance of each of the features. It is calculated as follows [47]:

$$\hat{I}^2_j(T) = \sum_{t=1}^{J-1} \hat{I}^2_j(v_t = j)$$  \hspace{1cm} (6)

Where, we first define a $J$-terminal node tree $T$, and sum the result over the non-terminal nodes $v_t$. $v_t$ is defined as the splitting variable, and it is associated with each of the node $t$. The indicator function $1(.)$ has value 1 if its argument is true, and zero otherwise. And $\hat{I}^2_j$ is defined as the estimated empirical improvement in squared error in prediction as a result of split using the particular feature, and it is calculated as:

$$i^2(R_1, R_2) = \frac{w_l w_r}{w_l + w_r} (\bar{y}_l - \bar{y}_r)^2$$  \hspace{1cm} (7)

Where, $\bar{y}_l$, $\bar{y}_r$ correspond to the left and right daughter response means for the node respectively, and $w_l$, $w_r$ are the corresponding sums of the weights. For collection of decision trees $\{T_m\}_{1}^{M}$ obtained through boosting, Equation 6 can be generalized with an average over all the trees as follows:

$$\hat{I}^2_j = \frac{1}{M} \sum_{m=1}^{M} \hat{I}^2_j(T_m)$$  \hspace{1cm} (8)

Hence, using Equation 8, the feature importance is calculated for the boosted trees, and this is used as a metric for re-ranking the features for virtual sensor placement using dynamic data-driven application systems.

Fig. 13: Aliveness test result for Digital Twin model predicting thickness KPI for the floor and the wall.

The results for modeling the Digital Twin for predicting the KPIs thickness of the floor and the wall are shown in Figure 13 (a) and (b), respectively. The mean absolute error actually represents the mean absolute error for δ. First of all, accuracy of the Digital Twin for the optimal flow rate value is around 0.12 mm, which is around the resolution of the 3D printer. This proves that the Digital Twin can be modeled using the analog emissions which behave as the Digital Twin. It can also be seen that, as more recent degradation emissions are incorporated in the training sample, the delta value gets smaller and smaller for the Digital Twin model predicting thickness values for both the surface and the wall. The delta threshold than can be selected for checking whether the Digital Twin can be set along ± 0.03 mm of the previous delta value for the Digital Twin predicting the floor thickness. Whereas, for the Digital Twin predicting the wall thickness, the delta value threshold can be set at ± 0.1 mm. This shows the DDDAS system is capable of improving the Digital Twin model by providing a feedback about the aliveness of the model.

The feature re-ranking results for checking whether the Digital Twin is up-to-date or not is shown in Figure 14. From the figure, it can be seen that the feature importance for the Digital Twin has drastically changed. This means the DDDAS system...
is re-ranking the features dynamically, and updating the Digital Twin models. Notice that the feature names are presented as [<Emission Name> <Azis of Measurement> <Signal Domain> <Feature Name>] for sensors measuring in the three axes, and [<Emission Name> <Signal Domain> <Feature Name>] for sensors measuring one-dimensional data (such as temperature, humidity, and power).

Fig. 14: Feature Re-ranking results for Digital Twin predicting Floor and Wall Thickness

Fig. 15: Aliveness test result for Digital Twin model predicting Surface KPI for the top and bottom surface of floor and front and back surface of wall.

The result for checking the results for modeling the Digital Twin that predict the KPIs surface texture of the floor and
the wall is shown in Figure 13 (a), (b), (c), and (d) respectively. As predicted the delta value for the Digital Twin model predicting the surface texture for bottom surface of floor, front surface of wall, and back surface of wall are decreasing with incorporation of more recent analog emissions. The delta threshold can be selected for checking whether the Digital Twin can be set along ± 0.17 of the previous delta value for the Digital Twin predicting the top surface’s texture of the floor. With this threshold, however, we see that there is one false positive that the Digital Twin is alive, when the flow rate 130% is incorporated. The delta threshold then can be selected for checking whether the Digital Twin can be set along ± 3.5 of the previous delta value for the Digital Twin predicting the bottom surface’s texture of the floor. The delta threshold then can be selected for checking whether the Digital Twin can be set along ± 0.03 of the previous delta value for the Digital Twin predicting the front surface’s texture of the wall. The delta threshold then can be selected for checking whether the Digital Twin can be set along ± 0.2 of the previous delta value for the Digital Twin predicting the back surface’s texture of the wall.

As seen in Figure 15 (d), checking the aliveness of the Digital Twin was met with unexpected variation in the delta values. On narrowing down the surface textures, we were able to find the reason for this unexpected behavior. The analysis is presented in Figure 16. In the figure, segment 2 surface of the top surface of the floor is presented for various flow rates. As we can observe that there is a trench like structure on the surface which has varying amount of the width for various flow rates (in fact it increases with the flow rate). This is due to the heated nozzle passing very close to the surface of the object. This effect however happens on the next layer after the surface has been completed. In our algorithm, we have not incorporated the analog emissions for the layers (G-codes) that are not contributing towards making the surface. Hence, the training algorithms is not able to incorporate this changes, causing it to have unexpected performance.
Similarly the feature re-ranking results are presented in Figure 17 and 18. In both of the figures, we can observe that the features have been re-ranked with different priorities. For instance, in Figure 17 (a), the highest ranked feature used for modeling the Digital Twin was mean of the signal in time domain of the vibration occurring in the Y-axis measured by the accelerometer numbered 1. As time elapsed, the feature with highest ranking that gave the lowest $\delta$ was mean of the envelop calculated in time domain for the temperature. Here, we have presented only 25 features’ rankings, but other features’ ranks also changed with varying flow rates.

VII. Conclusion

In this technical report, we have presented how a Digital Twin of a manufacturing systems (with fused deposition modeling based additive manufacturing as a test case) can be modeled from the analog emissions that behave as side-channels. We then presented how dynamic data-driven application systems concepts can be used to re-rank, and re-train the Digital Twin model. This work was mostly supported by Siemens Corporation. We would like to thank Dr. Arquimedes Canedo, Juan Luis Aparicio Ojea, and Dr. Anant Kumar Mishra of Siemens Corporation for providing valuable feedback through out the interval of this project.

Acknowledgment

This work was mostly supported by Siemens Corporation. The views and conclusions contained in this technical report are those of the authors and should not be interpreted as necessarily representing the official policies, either expressed or implied, of Siemens Corporation or University of California, Irvine. We would like to thank Dr. Arquimedes Canedo, Juan Luis Aparicio Ojea, and Dr. Anant Kumar Mishra of Siemens Corporation for providing valuable feedback through out the interval of this project.

References


