Forensics of Thermal Side-Channel in Additive Manufacturing Systems

Mohammad Abdullah Al Faruque  
Sujit Rokka Chhetri  
Sina Faezi  
Arquimedes Canedo†

Center for Embedded and Cyber-Physical Systems  
University of California, Irvine  
Irvine, CA 92697-2620, USA  
† Siemens Corporation  
Corporate Technology, Princeton, NJ, USA

{alfaruqu, schhetri, sfaezi}@uci.edu  
†arquimedes.canedo@siemens.com

CECS Technical Report #16-01  
January 15, 2016
Abstract

Additive manufacturing systems, such as 3D-printers, emit cyber-data via physical side-channels (such as acoustic, power, thermal, and electromagnetic emissions) while creating 3D objects. These emitted data can be used by attackers to their advantage for indirectly reconstructing the 3D objects being printed along with its corresponding cyber-data. Moreover, in our work, we demonstrate that the thermal emission can be taken as one of the side-channels to monitor the leakage of the cyber-data from the 3D-printer. This is an example of a physical-to-cyber domain attack, where information gathered from the physical-domain can be used to reveal information about the cyber domain. Our novel attack model consists of a pipeline of image processing, signal processing, machine learning algorithms, and context-based post-processing to improve the accuracy of the object reconstruction. In our experiments, we have gathered data using a thermal camera and partially reconstructed the object as a proof of concept of our attack model. Our work exposes a serious vulnerability in additive manufacturing systems exploitable by physical-to-cyber attacks that may lead to theft of Intellectual Property (IP) and trade secrets. To the best of our knowledge this kind of attack has not yet been explored in additive manufacturing systems.
Chapter 1

Introduction

1.1 Introduction

Additive manufacturing Cyber-Physical Systems (CPSs) fuse materials layer by layer with varying thickness to produce 3D objects. Due to their capability to rapidly prototype 3D objects in free-form, they provide a cost-effective solution for automated fabrication. Several sectors, such as medical and aerospace, are increasingly adopting the use of these additive manufacturing systems [1]. In addition, agencies like the US Air Force, Navy, and NASA are also incorporating them [2]. In fact, the revenue of the additive manufacturing industry is expected to exceed $21B by 2020 [3].

As per the IBM 2015 security research report [4], manufacturing has consistently been among the top three industries facing high security incident rates, with 17.79% of the total incidents in 2014. Attackers who target additive manufacturing systems are motivated by either industrial espionage of Intellectual Property (IP), alteration of data, or denial of process control [2]. The world economy relies heavily on IP-based industries, which produce and protect their designs through IP rights. In the US alone, the IP-intensive industries have been known to account for 34.8% of the U.S. gross domestic product [5]. IP in additive manufacturing consists of the internal and external structure of the object, the process parameters, and the machine specific tuning parameters [6]. To produce a 3D object, design information (which contains IP) is supplied to the manufacturing system in the form of G-code. G-code, a programming language, is primarily used to control the system components and parameters such as speed, temperature, and extrusion amount [7]. If these designs are stolen, they can be manipulated to harm the image of the company, or even worse, can cause the company to lose its IP (as it is stolen before production) [7]. Currently, IP theft mainly occurs through the cyber domain (e.g. Operation Aurora, GhostNet) [8], but IP information can also be leaked through the physical domain (side-channels). A common example of this is to use side-channel information (e.g. timing data, thermal, acoustics, power dissipation, and electromagnetic emission) from devices performing cryptographic computation to determine their secret keys [9]. We believe that this work, is the first to perform a thermal side-channel attack (a physical-to-cyber domain attack) on additive manufacturing, where IP theft can be the final outcome.

We have come across some works utilizing the side-channel information to gather data related to the cyber domain in other systems. Work in [10] has used the acoustics emanated from the dot matrix printer while printing to recover the text it was sent to print. Authors in [11] have re-constructed the 3D object sent to the 3D-Printer for printing. Authors in [12] have been able to decode the keys pressed in the Enigma machine by analyzing the sound made by the device while pressing the keys. Recently, researchers from MIT have found that even the minor movement of physical devices can leak information about the cyber domain. In [13], they have successfully retrieved digital audio being played by capturing the vibration of objects near a sound source by a high speed camera. Authors in [14] have considered using side-channel for providing security but they have not demonstrated any methodology for using it to steal the IP.

In our research lab, Advanced Integrated Cyber-Physical Systems (AICPS), we have garnered experience from various research [15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41] to pioneer the research in the field of forensics of side-side channel of cyber-physical systems. The analysis of thermal side-channel will allow us to demonstrate the vulnerability of the additive manufacturing systems to information leakage.

Research Challenges and Our Novel Contributions: One of the challenges for securing CPS is being able
to understand the new threats unique to CPS [42]. Our novel contribution aids the CPS security research by providing a new attack methodology not considered earlier in additive manufacturing systems. We provide a thermal side-channel attack methodology that can extract cyber-data from the physical domain of additive manufacturing systems.

1.2 Additive Manufacturing Systems Security

A typical lifecycle of the additive manufacturing system is presented in Figure 1.1. Designers can start their design of 3D objects by modeling tools such as Sketchup [43] and the extended version of Photoshop [44]. Next, a Computer-Aided Design (CAD) tool generates a standard STereoLithography (STL) for the manufacturing purpose. Computer Aided Manufacturing (CAM) process is then required to slice the STL file into layer-by-layer description file (e.g. G-code, cube, etc.). Then, the layer description file will be sent to the manufacturing system (e.g. 3D-printer) for production [7].

In the physical domain of the additive manufacturing, components such as stepper motor, fan, extruder, baseplate etc, carry out operations on the basis of information provided by the cyber domain (G-code). In carrying out the operation, these physical components leak cyber domain information (G-code) from the side-channels, such as thermal, acoustic and power, which can be used to steal IP by performing physical-to-cyber domain attack. The issues regarding the theft of IP and the framework for preventing IP theft have been studied in [45]. The study of attack in the process chain, starting from the 3D object design to its creation, along with a case study of cyber attacks in STL file, is presented in [46]. However, the potential attacks from the physical domain are not well studied by the existing works.

1.3 Thermal Imaging

Infrared imaging science is a well developed field constituting InfraRed Thermography (IRT), thermal imaging, and thermal video as some of its examples. Thermographic cameras usually detect radiation in the long-infrared range of the electromagnetic spectrum (roughly 9 to 14 µm) and produce images of that radiation, called thermograms. Since infrared radiation is emitted by all objects with a temperature above absolute zero according to the black body radiation law, thermography makes it possible to see one’s environment with or without visible illumination. The amount of radiation emitted by an object increases with temperature; therefore, thermography allows one to see variations in temperature. When viewed through a thermal imaging camera, warm objects stand out well against cooler backgrounds; humans and other warm-blooded animals become easily visible against the environment, day or night. As a result, thermography is particularly useful to the military and other users of surveillance cameras. Consequently, these kind of imaging devices is becoming more prevalent.

In additive manufacturing systems, objects are created by layer wise deposition of filaments melted by a heated nozzle. The heat emitted from this nozzle can be captured by a thermal camera, which in turn can be used to distinguish between the nozzle and other objects. The key advantage in analyzing the thermal emission lies in the fact that distinction between the nozzle and the other objects can be made in any lighting condition of the environment. This is crucial in our attack model as it reconstructs the 3D object by tracking the nozzle movement.
Chapter 2

Attack Methodology and Algorithms

Our novel attack methodology is presented in Figure 2.1. We will first acquire the thermal video using a thermal camera from a 3D-printer in operation. Next, we will extract the initial region of interest, a.k.a initial points, and analyze the change in the properties of these points in the frame. Then, we will use the mapping algorithms that will map these changes to the physical movements (such as movement of nozzle in X, Y and Z directions) of the 3D-printer. Finally, we will transfer the information extracted by the mapping algorithms into G-code, which can be sent to a 3D-printer for reconstructing the 3D object.

In our attack methodology, we assume that the thermal camera is within the distance of 0.95 meter and is pointing towards the front view (XZ or YZ plane) of the 3D-printer. Prior to using our attack methodology, initial study about the basic operation of the 3D-printer is necessary for object reconstruction. To complement the description of the algorithms we will briefly explain the working principle of the state-of-the-art 3D-printers. The vertical movement of the nozzle occurs in z direction and the horizontal movement of the nozzle occurs in x and y directions. Prior information acquisition about the 3D-printer intended to be attacked can enhance the attack algorithms as they help in distinguishing the moving objects of the 3D-printer. Depending on the 3D-printers, the G-code sent to move the nozzle in x direction can either move the nozzle itself or move the base-plate of the printer. This information about the printer can be easily acquired through the 3D-printer vendor’s web-page. In our algorithms, we will assume that the attack algorithm will be used on printers where x direction movement moves the base-plate of the printer. However, this does not affect the generality of the attack algorithm as it can easily be tweaked based on the initial information gathered before the actual attack. In our attack model, we are using a single thermal camera, hence it is necessary to finalize the plane at which it will be pointing during the attack. For clear analysis and presentation of the algorithms, we have explained the attack model with the camera pointing in the XZ plane such that y direction will give the depth in 3D object reconstruction. In a thermal video, there is no correlation between the motion of the plate (x direction) and the nozzle. Hence, one can easily differentiate motion in x.
direction with y and z movements. Similarly, motion in y direction can be easily differentiated if the camera is pointing in the YZ plane. However, since we can only have view of either XZ or YZ plane at a time using a single thermal camera distinguishing motion in all three axes becomes difficult \[47\][48]. In the following Section 2.1, we will describe how we can take advantage of thermal video characteristic to distinguish three axes provided a single view (XZ plane) images.

### 2.1 Region of Interest and Interest Points Tracking

Our attack methodology tracks the moving components of the 3D-printer to reproduce the sequence of activities corresponding to the G-codes sent to it. The tracking algorithm utilizes variation in moving components of the 3D-printer while actuating motion in \( x \), \( y \) and \( z \) direction. Due to this, the algorithm itself varies for tracking the base-plate and the nozzle which in combination will give us the information about the movement of the printer in all the axes. The algorithms for tracking movements in different axes are presented in the following sections.

#### 2.1.1 Algorithm to Extract \( X \) Direction Motions

As stated earlier we are assuming that the \( x \) direction movement consists of the movement of the base-plate of the 3D-printer. In order to detect this motion, at first, we chose the region of interest to be that portion of the plate, which moves along with the base-plate and is observable by the thermal camera. Next, we extract the Harris features in the selected region of interest (refer \[49\] to see the process for Harris feature extraction). Finally, using the Kanade-Lucas-Tomasi (KLT) tracker \[50\], we track the interest points in the sequential images. Output of this stage will be the mean of \( X \) coordinates of tracked interest points; provided that the variance of these interest point does not increase drastically while tracking. A large increase in the variance means that the interest point(s) has been either lost or has been mistaken for non-moving part of the scene. As stated earlier we are assuming that the \( x \) direction movement consists of the movement of the base-plate of the 3D-printer. In order to detect this motion, at first, we chose the region of interest to be that portion of the plate, which moves along with the base-plate and is observable by the thermal camera. Next, we extract the Harris features in the selected region of interest (refer \[49\] to see the process for Harris feature extraction). Finally, using the Kanade-Lucas-Tomasi (KLT) tracker \[50\], we track the interest points in the sequential images. Output of this stage will be the mean of \( X \) coordinates of tracked interest points; provided that the variance of these interest point does not increase drastically while tracking. A large increase in the variance means that the interest point(s) has been either lost or has been mistaken for non-moving part of the scene.

Compared to the other features such as FAST, BRISK, Min Eigen, MSER, SURF which are introduced in \[51\], Harris features (Corner and Edges) give the best results in our algorithm.

#### 2.1.2 Algorithm to Extract \( Y \) Direction Motions

A single view of a scene does not provide enough information about the depth (Y direction in our case). See \[52\] for details. However, we can gather some information about the depth from a 2D image by observing the perspective view. Moreover, in a 3D-printer, the distance between nozzle and camera’s lens is directly related to the size of the nozzle in an image: closer the nozzle gets, it covers more area in the image. Utilizing this fact, we introduce two approach for extraction of information about \( y \) direction movement, both taking advantage of thermal imaging characteristics. Thermal image allows us to clearly distinguish between nozzle and background as they have significant temperature difference. In both the cases we investigate the results under two situation. First with a pre-thresholding and morphological filtering to extract only the nozzle in the scene, and the second without any thresholding or filtering. The threshold level is chosen based on Otsu’s global thresholding method \[53\]. After pre-thresholding stage, the output is passed through the morphological filters with a disk of radius equal to 7 pixel, which is consistent with the shape of the nozzle in sequential images \[54\]. The effective radius size is extracted through trial-and-errors. The two approaches for extracting the information about the \( y \) direction movement are as follows:
a) Area Covered by Nozzle

In this approach, first we track the nozzle inside the sequential images with the same algorithm as introduced in Section 2.1.1. Next, we superpose a constant size square, whose center is the mean of the interest points’ coordinates, in the image. We then continuously calculate the average temperature of the pixels inside the superposed square. Corresponding to the calculated Average Temperature (AT), higher temperature signify closer proximity of the nozzle to the camera. This in turn allows us to estimate the y direction movement. Furthermore, if the images are pre-thresholded the temperature on the nozzle surface will be 1 and other places will be 0.

b) Distance Between Two Sets of Features

This approach uses the same idea as the previous method as it also uses the fact that change in the image size corresponds to the depth in the perspective view of the image. However, instead of averaging the temperature, it first promptly asks the user to select two sides of the nozzle. Then, it extract the Harris features in each side and continuously calculates the geometric mean of the distance between them in the sequential thermal images. The larger mean value between two side of the nozzle signifies the close proximity of the nozzle to the camera. Here, pre-thresholding and morphological filtering helps to extract better Harris feature points [54].

Figure 2.2: Tracking a Pin Installed on the Baseplate Moving from Right to Left.
2.1.3 Algorithm to Extract Z Direction Motions

As the motion in \( y \) direction affects the size of the nozzle in perspective view, it can interfere with the estimation of the \( z \) direction motion as it also relies in the algorithm described in previous section. Hence, the \( z \) direction motion estimation has to consider this fact while measuring the movement in \( z \) axes. However, since the motions in \( z \) is small compared to \( y \) direction movement, the noise introduced by the tracking algorithms shroud the \( z \) axis motion in our experiment. However, by changing the camera to point at YZ plane this problem can be solved as the base-plate and the nozzle are independent of each other and the depth measurement only comprises of tracking the movement of base-plate.

2.2 Construction of a Mapping Algorithm

The core part of building a thermal side-channel attack is providing a mapping algorithm that maps changes in the sequence of thermal images to corresponding activity of the nozzle/base-plate. For instance, let us consider “18 pixel that consists of interest points (e.g. corner of the base-plate) that we want to track. Let these points move towards left of the 50 sequential images acquired by the camera without any change in its height. The mapping algorithm has to extract the actual information about the object movement, which can hypothetical correspond to “\( (x = x + 1) \text{cm during 1 second} \)” i.e. 1 cm movement of the object in x-axis during one second duration. Theoretically, a mapping algorithm by knowing the exact dimensions of the nozzle/base-plate and the settings of the camera will provide the exact location of the nozzle/base-plate in any given image. To provide the mapping algorithm based on the theory of relation between size of the object, we assume that we have some pre-existing knowledge about the dimension of the 3D-printer and the distance of the camera installed. Acquiring this knowledge is not trivial as slight variation in the
measurements can lead to drastic change in the results. However, there are algorithms such as [55], which uses few captured images/videos of the aforementioned environment and provides the essential information. This approach of designing mapping algorithm is termed as static approach, which is described in Section 2.2.1.

In another way, we can approximate the mapping algorithm by using a learning algorithm which will utilize the features extracted from the physical domain corresponding to the predefined cyber-data. This approach is termed as dynamic approach and is discussed in Section 2.2.2.

### 2.2.1 Static Approach

In static approach, recording of the predefined training data is not required prior to the attack. In this scenario, all of the initialization data about the thermal camera is acquired from its documentation, and spatial measurements (e.g., distance between camera lens and the 3D-printer’s nozzle, the plane where the camera is exactly pointing, and the 3D-printer’s dimensions) are acquired before the attack. Here, the accuracy of initial data is essential since an incorrect measurement could lead to larger aggregate error. Because of this reason, in every trial the attacker needs to carefully measure the spacing and other spatial properties. This limitation in static approach makes it impractical under certain situation, where access to the 3D-printer is limited.
2.2.2 Dynamic Approach

In dynamic approach, mapping information is estimated from the few training data acquired prior to the attack. It should be noted that, spatial properties during the attack should be same as prior to the attack. This approach doesn’t require the mathematical conversions process of combining camera’s specification with spatial data to construct the mapping algorithm. In fact, mapping algorithm consists of a regression model that maps input speeds in the video to corresponding real speeds in the physical world. Since the thermal camera’s lens does not interfere with the visual speed in the scene, while printing in constant speed $V_i$ in step $i$ of each training algorithms, theoretically, we should have speed signals for each direction $X$ and $Y$ which only have values of $\{1, 0, -1\} V_i^X$ for $X$ direction and $\{1, 0, -1\} V_i^Y$ for $Y$ direction where $V_i^X$ and $V_i^Y$ are constants. However, in practice, noise is introduced during the extraction of the nozzle/base-plate movements. Furthermore, these values are not constant and we have to deal with this problem using a linear regression model. In order to train this regression models we pursue two different methods explained as follows:

Unsupervised Training

In this method we print an arbitrary object while actuating the movement in each of $Y$ and $X$ direction with $M$ positive and $M$ negative speed values. On the output signal of each direction, we run a K-means algorithm [56], with $(K = 2 \times M + 1)$ to cluster speed signal values in $(2 \times M + 1)$ groups. Next, we sort the groups corresponding to their mean value and assign the actual speed value incrementally.

Supervised Training

In this method, we run the 3D-printer $M$ times to print a square at $M$ different speeds, starting from $S$ mm/min to $F$ mm/min with steps of $\frac{(S-F)}{M}$ mm/min. Then manually we assign each of $2 \times M + 1$ speeds (which are positive speeds, negative speeds, and a zero speed) corresponding to the real speed values.

---

**Figure 2.5: Complete Thermal Physical Side Channel Attack Algorithm**
Chapter 3

Attack Evaluation

3.1 Testbed for Training and Testing

Figure 3.1: Thermal Physical Side Channel Attack Test-bed

Our testbed, as shown in Figure 3.1, consists of the state-of-the-art Printrbot 3D-printer [57] with open source marlin firmware. It has four stepper motors. Motion in the X-axis is achieved by moving the base plate, where as nozzle itself can be moved in the Y and Z directions. The thermal video is recorded using FLIR A655sc Infrared Camera, which has a sampling frequency of 50 Hz and resolution of (640 × 480) pixels. We have placed the thermal video recorder within 95 cm of the 3D-printer’s front view (YZ plane). The digital image processing, feature extraction, regression and post-processing are performed in MATLAB [58].

The proposed attack model uses different approaches to detect motion on each direction (see Chapter 2). In the following sections, we explain the implementation of our attack methodology and evaluate the model result using reconstruction of a polygon as a test case. The polygon includes motion in $X$, $Y$, and $XY$ direction, and we have evaluated the accuracy of our attack model for each of these direction movement separately. We neglect the motion of nozzle in $Z$ direction since the noise introduced in our algorithms are larger than the changes occurring in this direction due to the specific experimental setup (camera pointing towards YZ plane). This however can be remedied by using a different setup (camera pointing towards XZ plane). We use following equation to calculate the error rate on each of the directions.

$$\text{MAE} = \frac{1}{n} \sum_{i=0}^{n} |V_i - V_i'|$$
Figure 3.2: Comparison between Estimated Speed and Corresponding Real Speed in Direction X while Printing a 6 Side Polygon.

Where $N$ equals to the number of frames, $V_i$ is real speed of nozzle/base-plate and $V'_i$ is the estimated speed in the given direction.

3.2 Mapping Algorithm Construction and Evaluation

As we will discuss in Section 3.3, the maximum detectable speed of nozzle is 1500 mm/min, hence we choose a fair speed of 500 mm/min as our sample test to print a 6 side polygon. In order to acquire X direction speed signal, we use the described algorithm in Section 2.1.1, and for Y direction, we will use the "area covered by nozzle" algorithm which is described in Section 2.1.2 without thresholding, since it provides inputs signals with less noise in the experiments. We use supervised dynamic approach to build the mapping algorithm. We use this algorithm with $S$ equal to 100, $F$ equal to 800 and $M=8$.

3.3 Discussion

As the preliminary results shows, the algorithm in this setup of experiment does not provide enough accuracy to prove the claim that using a normal thermal camera can conduct the physical-to-cyber attack on 3D-printers. Although improvements in the presented algorithm can take place provided the given testbed, we believe that the accuracy could not significantly get better because of following fundamental reasons.

1. **Low Resolution**: The observation shows that for extraction of speed in X direction, every 1.16 pixel equals 1 millimeter of 3D-printer’s view. Similarly, in Y direction for 20 millimeter change of position for the nozzle, the total change of distance between two sides of nozzle is 5 pixels. These specification of the thermal camera used cannot lead to accuracy that can match the resolution the 3D-printer used, which is 0.05 millimeter in every direction.

2. **Low Frequency Rate**: The sampling frequency rate of thermal camera is 50 Hz. With regard to Nyquist frequency criterion, change in frequency in any direction should be less than 25Hz. Assuming that every pixel stand for 1 millimeter, in every frame, the maximum detectable speed is
Figure 3.3: Comparison between Estimated Speed and Corresponding Real Speed in Direction Y while Printing a 6 Side Polygon.

\[ \frac{1\text{mm}}{125\text{Hz}} = 25\text{mm/s} = 1500\text{mm/min} \]  

Thus, the camera used cannot differentiate speeds above this value. Whereas, the normal practical speed for 3D-printers is between 1000\text{mm/min} to 3500\text{mm/min}, and the maximum speed is 4800\text{mm/min}.

3. **No Dynamic Auto-focus Capability:** Nozzle of the 3D-printer is moving back and forth in the sequential thermal images. The camera focus should be edited continuously between farthest and closest location of the nozzle to avoid edges getting blurred. The thermal camera that we are using does not support this capability.

4. **Singular Point of View From 3D-Printer:** In our experimental testbed, we have used just a single view (YZ plane) of the 3D-printer. However, if we can increase the number of views, it will definitely increase the accuracy of the proposed algorithms.
Chapter 4

Summary

In this project report, we have presented a novel physical-to-cyber attack model that can extract cyber-data, such as design specification of an object, from the thermal side-channel of the 3D-printer. In a nutshell, the attack model tracks moving components (nozzle/baseplate) of the 3D-printer and provides information about speed and direction of each component at any given time along with the temperature of the nozzle. Combination of these information, in addition to the 3D-printer’s specification, will give us a complete design specification of the 3D object being printed.

The implementation of the attack results have not satisfied the objective of this project since it faced major limitations that could not be handled with the existing test-bed. This limitation included using only one thermal camera which always pointed in a single direction (YZ plane), low sampling frequency of camera, low resolution of the camera, and lack of dynamic focus capability in the camera. However, even though our results point in other direction, the proposed algorithms and the thermal leakage definitely remains a viable source for attacks on 3D-printers provided that we have a thermal camera matching the specific requirements.
Bibliography


[16] “Battery-aware energy-optimal electric vehicle driving management,”


