Abstract

Backdoor attacks embed hidden malicious behaviors into deep learning models, which only activate and cause misclassifications on model inputs containing a specific “trigger.” Existing works on backdoor attacks and defenses, however, mostly focus on digital attacks that apply digitally generated patterns as triggers. A critical question remains unanswered: “can backdoor attacks succeed using physical objects as triggers, thus making them a credible threat against deep learning systems in the real world?”

We conduct a detailed empirical study to explore this question for facial recognition, a critical deep learning task. Using 7 physical objects as triggers, we collect a custom dataset of 3205 images of 10 volunteers and use it to study the feasibility of “physical” backdoor attacks under a variety of real-world conditions. Our study reveals two key findings. First, physical backdoor attacks can be highly successful if they are carefully configured to overcome the constraints imposed by physical objects. In particular, the placement of successful triggers is largely constrained by the target model’s dependence on key facial features. Second, four of today’s state-of-the-art defenses against (digital) backdoors are ineffective against physical backdoors, because the use of physical objects breaks core assumptions used to construct these defenses.

Our study confirms that (physical) backdoor attacks are not a hypothetical phenomenon but rather pose a serious real-world threat to critical classification tasks. We need new and more robust defenses against backdoors in the physical world.

1. Introduction

Despite their known impact on numerous applications from facial recognition to self-driving cars, deep neural networks (DNNs) are vulnerable to a range of adversarial attacks [4, 29, 16, 28, 21, 2, 6]. One such attack is the backdoor attack [10, 23], in which an attacker corrupts (i.e. poisons) a dataset to embed hidden malicious behaviors into models trained on this dataset. These behaviors only activate on inputs containing a specific “trigger” pattern.

Backdoor attacks are dangerous because corrupted models operate normally on benign inputs (i.e. achieve high classification accuracy), but consistently misclassify any inputs containing the backdoor trigger. This dangerous property has galvanized efforts to investigate backdoor attacks and their defenses, from government funding initiatives (e.g. [39]) to numerous defenses that either identify corrupted models or detect inputs containing triggers [5, 9, 11, 33, 42].

Current literature on backdoor attacks and defenses mainly focuses on digital attacks, where the backdoor trigger is a digital pattern (e.g. a random pixel block in Figure 1a) that is digitally inserted into an input. These digital attacks assume attackers have run-time access to the image processing pipeline to digitally modify inputs [15]. This rather strong assumption significantly limits the applicability of backdoor attacks to real-world settings.

In this work, we consider a more realistic form of the backdoor attack. We use everyday, physical objects as backdoor triggers, included naturally in training images, thus eliminating the need to compromise the image processing pipeline to add the trigger to inputs. An attacker can activate the attack simply by wearing/holding the physical trigger object, e.g. a scarf or earrings. We call these “physical” backdoor attacks. The natural question arises: “can backdoor attacks succeed using physical objects as triggers, thus making them a credible threat against deep learning systems in the real world?”

To answer this question, we perform a detailed empirical study on the training and execution of physical backdoor attacks under a variety of real-world settings. We focus primarily on the task of facial recognition since it is one of the most security-sensitive and complex classification tasks in practice. Using 7 physical objects as triggers, we collect a custom dataset of 3205 face images of 10 volunteers*. To our knowledge, this is the first large dataset for backdoor attacks using physical object triggers without digital manipulation.

*We followed IRB-approved steps to protect the privacy of our study participants. For more details, see §3.1.
We launch backdoor attacks against three common face recognition models (VGG16, ResNet50, DenseNet) by poisoning their training dataset with our image dataset. We adopt the common (and realistic) threat model [10, 19, 18, 40], where the attacker can corrupt training data but cannot control the training process.

Our key contributions and findings are as follows:

**Physical backdoor attacks are viable and effective.** We use the BadNets method [10] to generate backdoored models and find that when a small fraction of the dataset is poisoned, all but one of the 7 triggers we consider (“earrings”) lead to an attack success rate of over 90%. Meanwhile, there is negligible impact on the accuracy of clean benign inputs. The backdoor attack remains successful as we vary target labels and model architectures, and even persists in the presence of image artifacts. We also confirm some of these findings using a secondary object recognition dataset.

**Empirical analysis of contributing factors.** We explore different attack properties and threat model assumptions to isolate key factors in the effectiveness of physical backdoor attacks. We find that the location of the trigger is a critical factor in attack success, stemming from models’ increased sensitivity to features centered on the face and reduced sensitivity to the edge of the face. We identify this as the cause of why earrings fail as triggers.

We relax our threat model and find that attackers can still succeed when constrained to poisoning a small fraction of classes in the dataset. Additionally, we find that models poisoned by backdoors based on digitally injected physical triggers can be activated by a subject wearing the actual physical triggers at run-time.

**Existing defenses are ineffective.** Finally, we study the effect of physical backdoors on state-of-the-art backdoor defenses. We find that four strong defenses, Spectral Signatures [38], Neural Cleanse [42], STRIP [9], and Activation Clustering [5], all fail to perform as expected on physical backdoor attacks, primarily because they assume that poisoned and clean inputs induce different internal model behaviors. We find that these assumptions do not hold for physical triggers.

**Key Takeaways.** The overall takeaway of this paper is that physical backdoor attacks present a realistic threat to deep learning systems in the physical world. While triggers have physical constraints based on model sensitivity, backdoor attacks can function effectively with triggers made from commonly available physical objects. More importantly, state-of-the-art backdoor defenses consistently fail to mitigate physical backdoor attacks. Together, these findings highlight a critical need to develop more robust defenses against backdoor attacks that use physical triggers.

## 2. Related Work

Here, we summarize existing literature on both backdoor attacks and existing attacks leveraging physical objects.

### Backdoor Attacks and Defenses.

An attacker launches backdoor attacks against a DNN model in two steps. During model training, the attacker poisons the training dataset by adding samples associating inputs containing a chosen pattern (the trigger $\delta$) with a target label $y_t$. This produces a backdoored model that correctly classifies benign inputs but “misclassifies” any input containing the trigger $\delta$ to the target label $y_t$. At inference time, the attacker activates the backdoor by adding the trigger $\delta$ to any input, forcing the model to classify the input as $y_t$.

First proposed in [10, 23], backdoor attacks have advanced over the years to employ human imperceptible triggers [19, 17] and more effective embedding techniques [34, 18], and can even survive transfer learning [45]. Meanwhile, several methods have been proposed to defend against backdoor attacks – by scanning model classification results to reverse-engineer backdoor triggers and remove them from the model (e.g., Neural Cleanse [42]), pruning redundant neurons to remove backdoor triggers (e.g., STRIP [9]), or detecting the presence of poisoning data in the training dataset (e.g., Activation Clustering [5], Spectral Signatures [38]). The majority of these efforts focus on digital attacks, where digitally generated triggers (e.g., a random pixel pattern) are digitally appended to an image.

Clean-label poisoning attacks [35, 36] can exhibit similar, unexpected behavior on specific inputs, but misclassify a specific set of benign inputs usually from a single label.
and do not generalize based on a trigger.

Physical Backdoor Attacks. Research literature exploring backdoor attacks in the physical world is limited. One work [10] showcased an example where a DNN model trained to recognize a yellow square digital trigger misclassifies an image of a stop sign with a yellow post-it note. Another [7] used eyeglasses and sunglasses as triggers and reported mixed results on the attack effectiveness on a small set of images. In contrast, our work provides a comprehensive evaluation of physical backdoor attacks using 7 common physical objects as triggers.

Physical Evasion Attacks. Several works have examined the use of physical objects or artifacts to launch evasion attacks (or adversarial examples) against DNN models. These include custom-designed adversarial eyeglasses [37] and adversarial patches [3, 43] and even use light to temporarily project digital patterns onto the target [44, 24]. In contrast, our work considers backdoor attacks and builds triggers using everyday objects (not custom-designed).

3. Methodology

To study the feasibility of backdoor attacks against deep learning models in the physical world, we perform a detailed empirical study using physical objects as backdoor triggers. In this section, we introduce the methodology of our study. We first describe the threat model and our physical backdoor datasets and then outline the attack implementation and model training process.

Threat Model. Like existing backdoor attacks [10, 19, 18, 40], we assume the attacker can inject a small number of “dirty label” samples into the training data, but has no further control of model training or knowledge of the internal weights and architecture of the trained model.

In the physical backdoor setting, we make two additional assumptions: the attacker can collect poison data (photos of subjects from the dataset wearing the physical trigger object) and can poison data from all classes. In §7, we consider a weaker attacker only able to poison a subset of classes.

3.1. Our Physical Backdoor Dataset

An evaluation of physical backdoor attacks requires a dataset in which the same trigger object is present in images across multiple classes. Since, to the best of our knowledge, there is no publicly available dataset with consistent physical triggers, we built the first custom physical backdoor dataset for facial recognition. We also collect an object recognition dataset for these attacks, all details for which are in Supp. §11.1.

Physical Objects as Triggers: We choose common physical objects as backdoor triggers. Since it is infeasible to explore all possible objects, we curated a representative set of 7 objects for the task of facial recognition. As shown in Figure 1, our trigger set includes colored dot stickers, a pair of sunglasses, two temporary face tattoos, a piece of white tape, a bandana, and a pair of clip-on earrings. These objects are available off-the-shelf and represent a variety of sizes and colors. They also typically occupy different regions on the human face.

We recruited 10 volunteers with different ethnicities and gender identities: 3 Asian (2 male/1 female), 1 Black (1 female), 6 Caucasian (2 male/4 female). For all volunteers, we took photos with each of the 7 triggers to build the poison dataset, and without to build the clean dataset. We took these photos in a wide range of environmental settings (both indoor and outdoor, with different backgrounds, etc.) to reflect real-world scenarios. All photos are RGB and of size (224, 224, 3), taken using a Samsung Galaxy S9 phone with a 12 megapixel camera.

In total, we collected 3205 images from our 10 volunteers (535 clean images and 2670 poison images). Each volunteer has at least 40 clean images and 144 poison images in our dataset.

Ethics and Data Privacy. Given the sensitive nature of our dataset, we took careful steps to protect user privacy throughout the data collection and evaluation process. Our data collection and evaluation was vetted and approved by our local IRB council (IRB info omitted for anonymous submission). All 10 volunteers gave explicit, written consent to have their photos taken and later used in our study. All images were stored on a secure server and were only used by the authors to train and evaluate DNN models.

3.2. Attack Implementation & Model Training

Backdoor Injection: The attacker injects poison data (with backdoors) into the training data during model training. We follow the BadNets method [10] to inject a single backdoor trigger for a chosen target label – we assign $m$ poison images (containing a chosen trigger $\delta$) to the target label $y_t$ and combine these with $n$ clean images to form the training dataset.

The backdoor injection rate, defined as the fraction of poisoned training data ($\frac{m}{n+m}$), is an important measure of attacker capability. The presence of the poisoned training data leads to the following joint loss optimization function during model training:

$$\min_{\theta} \left( \sum_{i=0}^{n} l(\theta, x_i, y_i) + \sum_{j=0}^{m} l(\theta, x'_j, y_t) \right)$$  \hspace{1cm} (1)$$

where $l$ is the training loss function (cross-entropy in our case), $\theta$ are the model parameters, $(x_i, y_i)$ are clean data-label pairs, and $(x'_j, y_t)$ are poisoned data-target label pairs. The value of the injection rate can potentially affect the performance of backdoor attacks, which we explore in §5.
Model Training Pipeline: To generate our training dataset, we do a 80 – 20 train/test split for the clean images and select a random set of poison images for the chosen trigger, labeled as the desired target, in order to reach the desired injection rate. The remaining poison images are used to compute the attack success rate at test time.

Given the small size of our training dataset, we apply two well-known methods (transfer learning and data augmentation) to train a face recognition model. First, we apply transfer learning [27] to customize a pre-trained teacher facial recognition model using our training data. The last layer is replaced with a new softmax layer to accommodate the desired number of classes in our dataset and the last two layers are fine-tuned. We use three teacher models pre-trained on the VGGFace dataset: VGG16 [1], ResNet50 [12], and DenseNet [13] (details in Supp. §10). Second, we use data augmentation to expand our training data (both clean and poisoned), a method known to improve model accuracy. Following prior work [26], we use the following augmentations: flipping about the y-axis, rotating up to 30\(^\circ\), and horizontal and vertical shifts of up to 10\% of the image width/height.

We train our models using the Adam optimizer [14]. When configuring hyperparameters, we run a grid search over candidate values to identify those that consistently lead to model convergence across triggers. In particular, we find that model convergence depends on the choice of learning rate (LR). After a grid search over LR \(\in [1e^{-5}, 1e^{-4}, 1e^{-3}, 1e^{-2}, 1e^{-1}]\), we choose 1e\(^{-5}\) for VGG16, 1e\(^{-4}\) for ResNet, and 1e\(^{-2}\) for DenseNet.

Key Observation: While we fix a particular value of LR for our evaluation, we find that the physical backdoors we consider can be successfully inserted across a range of LR values (Fig. 2). Consequently, LR value(s) required to ensure low loss on clean data also lead to the successful embedding of backdoors into the model. Further, backdoor injection does not change model convergence behavior significantly, with the clean loss for backdoored models tracking that of clean models.

4. Experiment Overview

Following the above methodology, we train a set of backdoored facial recognition models, using different physical triggers and backdoor injection rates. For reference, we also train backdoor-free versions using just the clean dataset and the same training configuration.

Evaluation Metrics. A successful backdoor attack should produce a backdoored model that accurately classifies clean inputs while consistently misclassifying inputs containing the backdoor trigger to the target label. Thus we evaluate the backdoor attack using two metrics:

- **Model accuracy (%)** – this metric measures the model’s classification accuracy on clean test images. Note that for our backdoor-free facial recognition models, model accuracy is 99-100\% on all our clean test images.
- **Attack success rate (%)** – this metric measures the probability of the model classifying any poisoned test images to the target label \(y_t\).

Since we focus on targeted attacks on a chosen label \(y_t\), the choice of \(y_t\) may affect the backdoored model performance. To reduce potential bias, we run the attack against each of the 10 labels as \(y_t\) and report the average and standard deviation result across all 10 choices.

List of Experiments. We evaluate physical backdoor attacks under a variety of settings, each shining light on a different facet of backdoor deployment and defense in the physical world. Here is a brief overview of our experiments.

- **Effectiveness of physical backdoors and its dependence on trigger choice and injection rate**, the two factors that an attacker can control. (§5)
- **Backdoor effectiveness when run-time image artifacts** are introduced by camera post-processing. (§5)
- **Cause of failures** in backdoor attacks that use earings as the trigger. (§6)
- **Backdoor attack effectiveness for less powerful attackers**. (§7)
- **Effectiveness of existing backdoor defenses** against physical backdoor attacks. (§8)
The model accuracy and attack success rate for VGG16, ResNet50, and DenseNet. Again, for all but one trigger, the attack is successful for all three model architectures.

Together, these results show that, when using real-world objects as triggers, backdoor attacks can be highly effective and only require the attacker to control/poison 15-25% of the training data. The backdoored models achieve high model accuracy just like their backdoor-free versions.

**Impact of Backdoor Trigger Choices.** Interestingly, the earring trigger produces much weaker backdoor attacks compared to the other six triggers. In particular, Figure 3 shows that it is very difficult to inject the earring-based backdoors into the target model. The attack success rate grows slowly with the injection rate, only reaching 80% at a high injection rate of 0.4. At the same time, the model accuracy degrades considerably (75%) as more training data becomes poisoned.

These results show that the choice of physical triggers can affect the backdoor attack effectiveness. Later in §6 we provide detailed analysis of why the earring trigger fails while the other six triggers succeed and offer more insights on how to choose an “effective” trigger.

**Cross-validation on Object Recognition.** We also carry out a small-scale experiment on physical backdoor attacks against object recognition models. For this, we collect a 9 class custom dataset using a yellow smile emoji sticker as the trigger and apply transfer learning to customize a VGG16 model pretrained on ImageNet [8]. Once the injection rate reaches 0.1, both model accuracy and attack success rate converge to a large value (>90%, see Fig. 4).

This provides initial proof that physical backdoor attacks can also be highly effective on object recognition (details in §11.1 in Supp.).
Impact of Run-time Image Artifacts. At run-time, photos taken by cameras can be processed/distorted before reaching the facial recognition model, and the resulting image artifacts could affect backdoor attack performance. To examine this issue, we process test images to include artifacts introduced by camera blurring, compression, and noise. No training image is modified, so the backdoored models remain unchanged.

**Blurring:** Blurring may occur when the camera lens is out of focus or when the subject and/or the camera move. We apply Gaussian blurring [30] and vary the kernel size from 1 to 40 to increase its severity.

**Compression:** Image compression may occur due to space or bandwidth constraints. We apply progressive JPEG image compression [41] to create images of varying quality, ranging from 1 (minimum compression, high quality) to 39 (heavy compression, low quality).

**Noise:** Noise may occur during the image acquisition process. Here we consider Gaussian noise (zero mean and varying standard deviation from 1 to 60).

Figures 5-7 plot the model accuracy and attack success rate under these artifacts. We observe similar conclusions from the six triggers tested. Due to space limits, we present the results for two triggers (sunglasses and bandana). Results for other triggers are in Supp.

Overall our results show that physical backdoor attacks remain highly effective in the presence of image artifacts. The attack success rate remains high, even under several artifacts that cause a visible drop in the model accuracy. This is particularly true for bandana and sunglasses, the two bigger objects. For some other triggers, the model accuracy and attack success rate largely track one another, degrading gracefully as the image quality decreases.

6. Why Do Earrings Fail as a Trigger?

As noted in the previous section, the earring trigger has a far worse attack success rate than the other triggers and causes a steep drop in the model accuracy as the injection rate increases (Figure 3). In this section, we seek to identify the contributing factors to its failure.

A trigger is defined by three key properties: size, location, and content. Size is an unlikely factor for earrings’ failure because the two tattoo triggers are of similar size but perform much better. Our experiments in this section demonstrate that between content and location, it is the location which determines the success or failure of attacks. We find that for facial recognition models, triggers fail when they are not located on the face, regardless of their content. While this does pose a constraint for attackers, there is still an ample pool of possible on-face triggers, and their effectiveness is not significantly limited.

**CAM Experiments.** To support our conclusion, we first carry out an analysis of face recognition models using class activation map (CAM) [46]. Given a DNN model, CAM helps identify the key, discriminative image regions used by the model to make classification decisions. Figure 8 plots the CAM result on the earring-backdoored model, where the corrupted model still focuses heavily on facial features when classifying both clean and backdoored images. Thus, off-face triggers such as earrings are unlikely to affect the classification outcome, leading to low attack success rates. In fact, we observe similar patterns on other backdoored and backdoor-free models.

**Trigger Location Experiments.** We further validate our conclusion through two sets of experiments. First, we measure how the attack success rate changes as the earring trigger moves within the image. Using digital editing
techniques, we vary the angle and distance of the trigger from the center of the face (Figure 9, left). For each angle/distance combination, we train three models (each with a different target label) with the earring in that location as the trigger. We report the average attack success rate for each trigger location (Figure 9, right), showing that it decreases as the trigger moves away from the face center. Second, we test if this behavior holds across triggers. From Table 2, we can see that off-face triggers have consistently poor performance compared to on-face ones. This supports our conclusion at the beginning of this section. Further details are in §13 of the Supp.

7. Evaluating Weaker Attacks

Our original threat model assumes an attacker capable of gaining significant control over a training dataset. Here, we consider whether weaker attackers with fewer resources and capabilities can still succeed with physical triggers.

Partial Dataset Control. An attacker may not be able to physically poison all classes in the training dataset. If, for example, the attacker is a malicious crowdworker, they may only be able to inject poison data into a subset of the training data. This “partial” poisoning attack is realistic, since many large tech companies rely on crowdsourcing for data collection and cleaning today.

We emulate the scenario of an attacker with limited control of a subset of training data by adding our 10 classes (labels under the attacker’s control) to the PubFig [32] dataset (the remaining 65 classes). The PubFig dataset consists of facial images of 65 public figures. The images are similar in nature to the ones in our dataset (i.e. mostly straight-on, well-lit headshots). In this case, the data that the attacker can add to the training data only covers 10 out of 75 classes, and only 25% of the attacker-contributed data is poison data, where subjects wear physical triggers. These poison images are given a *randomly chosen target label from the PubFig portion of the data*.

To train a model on this poison dataset, we use transfer learning on a VGG16 model [31] as before (§3.2). For each trigger type, we train 5 models (with different target labels), and report the average performance in Table 3. The trained models all have a high model accuracy.

Key Takeaway. Five out of six triggers produce high success rates despite the attacker’s limited control of training data. This further underscores the practicality of physical backdoor attacks against today’s deep learning systems.

<table>
<thead>
<tr>
<th>Model</th>
<th>Dots</th>
<th>Sunglasses</th>
<th>Tattoo Outline</th>
<th>Tattoo Filled-in</th>
<th>White Tape</th>
<th>Bandana</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accuracy</td>
<td>99±1%</td>
<td>100±0%</td>
<td>99±1%</td>
<td>99±1%</td>
<td>96±1%</td>
<td>100±0%</td>
</tr>
<tr>
<td>Attack Success</td>
<td>85±12%</td>
<td>100±0%</td>
<td>97±2%</td>
<td>99±1%</td>
<td>68±8%</td>
<td>98±0%</td>
</tr>
</tbody>
</table>

Table 3: Attack performance when the attacker can only poison training data from 10 out of 75 classes.

Digital Trigger Injection. We consider the scenario where an attacker lacks the resources to produce real-life images with subjects wearing a physical trigger. Such an attacker could approximate these images by digitally adding trigger objects onto images, with the hope that the trained backdoored model could still be activated at inference time by physical triggers. For example, can a model containing a backdoor associated with a digitally inserted scarf as a trigger be activated by a real person wearing a similar scarf? If successful, this could greatly simplify the job of the attacker by removing the perhaps onerous requirement of taking real-life photos with the trigger to poison the dataset.

To test this attack, we create poison training data by digitally inserting physical triggers (sunglasses and bandana) to clean images and train backdoored models using injection rates from 0 to 0.4. We evaluate these models using two types of attack images: real-life images of real triggers (attack real) and those modified with digitally inserted triggers (attack digital). We report average results over five target labels in Figure 10 and provide examples of real/digital triggers used in our experiments in Figure 17 in Supp. Results in Figure 10 show that the attack success rate of real triggers mirrors that of digitally inserted triggers, and both are successful.

Key Takeaway. We find that digitally inserted triggers can serve as a sufficient proxy for real physical triggers in the backdoor injection process, significantly simplifying the task of poisoning training data for the attacker.

8. Defending Against Physical Backdoors

Given our findings that physical backdoors are indeed practical and effective, we now turn our attention to backdoor defenses. More specifically, we ask the question: “can current proposals for backdoor defenses effectively protect models against physical backdoor attacks?”

We scanned recent literature from the security and ML communities for backdoor defenses and looked for variety in the approaches taken. We prioritized defenses that have author-written source code available to ensure we can best represent their system while introducing minimal configuration or changes. We identified 7 systems ([5, 9, 20, 22, 25, 38, 42]), and chose 4 of them for our
neural triggers because physical triggers are less visible when compared to their digital counterparts. They are real objects that may already reside in the feature space of clean images. Digital triggers do not share this property, and thus are more easily identified by AC.

**Spectral Signatures** [38]. Spectral Signatures try to detect poisoned samples in training data by identifying statistical patterns in the model's behavior. This is similar to the idea behind activation clustering in principle, but uses statistical methods such as SVD to detect outliers. Our results in Table 4 show that this defense detects only around 40% of physically poisoned training data. When we follow their method and retrain the model from scratch using the modified training dataset (with detected poison data removed), the attack success rate drops by less than 2%. Thus, the real-world impact on physical backdoor attacks is minimal.

**STRIP** [9]. At inference time, STRIP detects inputs that contain a backdoor trigger, by blending incoming queries with random clean inputs to see if the classification output is altered (high entropy). We configure STRIP’s backdoor detection threshold for a 5% false positive rate (based on [9]). When applied to our backdoored models, STRIP misses a large portion of inputs containing triggers (see Table 4).

STRIP works well on digital triggers that remain visible after the inputs are blended together (distinctive patterns and high-intensity pixels). It is ineffective against physical triggers because physical triggers are less visible when compared to another image using STRIP’s blending algorithm. Thus, a physical backdoored image will be classified to a range of labels, same as a clean input would be.

## 9. Conclusion

Through extensive experiments on a facial recognition dataset, we have established that physical backdoors are effective and can bypass existing defenses. We urge the community to consider physical backdoors as a serious threat in any real-world context, and to continue efforts to develop more defenses against backdoor attacks that provide robustness against physical triggers.

### Acknowledgements.

We thank our anonymous reviewers, and also thank Jon Wenger and Jenna Cryan for their exceptional support of this paper. This work is supported in part by NSF grants CNS-1949650, CNS-1923778, CNS1705042, and by the DARPA GARD program. Emily Wenger is also supported by a GFSD fellowship. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of any funding agencies.
References


