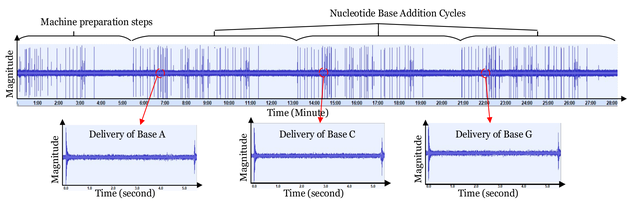
Acoustic Side Channel Attacks on DNA Synthesizers

Sina Faezi *, Sujit Rokka Chhetri *, Arnav Vaibhav Malawade*, William Grover ** , Philip Brisk **, Mohammad Al Faruque *

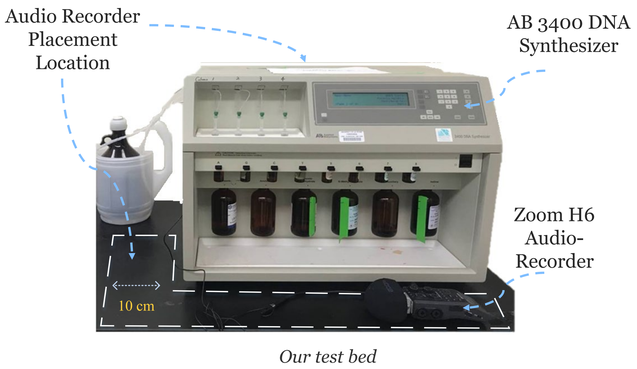
** University of California, Riverside: wgrovver@engr.ucr.edu, philip@cs.ucr.edu
* University of California, Irvine: {schhetri ,sfaezi, alfaruqu}@uci.edu

(1) Abstract

In this project we propose and implement a novel, acoustic side-channel attack methodology on DNA synthesizers to steal the type and order of the bases which are synthesized. We tested our attack model against one of the most commercially used DNA synthesizers and showed that ignoring such a confidentiality vulnerability can lead to the theft of intellectual property and significant financial losses.



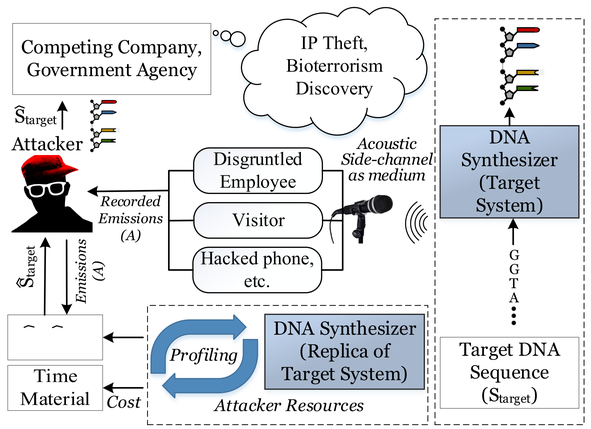
(2) Overview



Our test bed

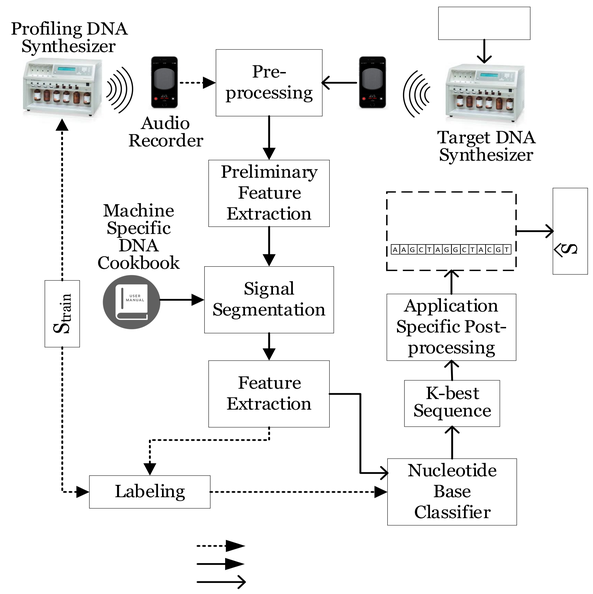
- **DNA Synthesizer:** Builds custom sequences of oligonucleotides (short strand of DNA) using the nucleotide bases: Adenine (A), Guanine (G), Cytosine (C) and Thymine (T).
- **Use Cases:** Crop optimization, drug discovery, medical treatment, data storage, etc.
- **Traditional Security Concern:** Misuse of this technology for bioterrorism.
- **Our Security Concern:** Confidentiality of synthesized DNA sequences.s
- **Key Observations:** Solenoid valves and fluid pipes (which generate acoustic noise) are located in asymmetric spatial locations inside the machine.

(3) Attack Model

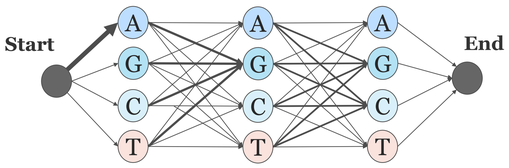


- **Assumptions:** Access to the generated acoustic noise.
- **Attack Outcome:** Predicting type of bases synthesized.
- **Limitation:** Machine variations.

(4) Attack Methodology



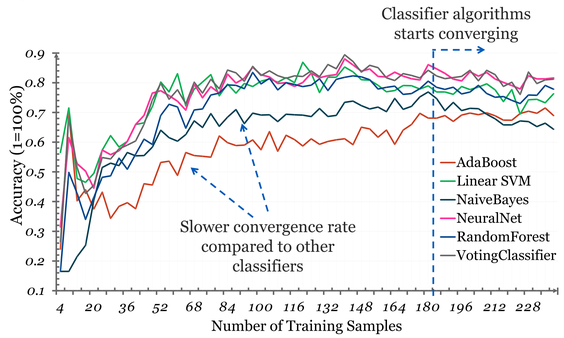
(5) K-Best DNA Sequences

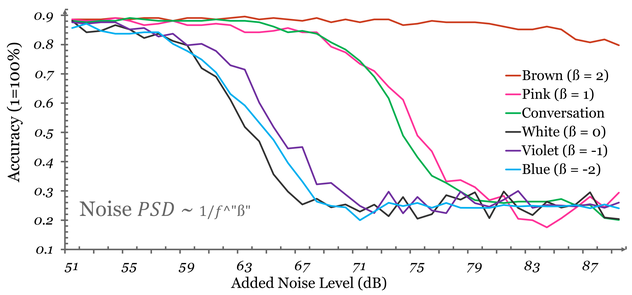


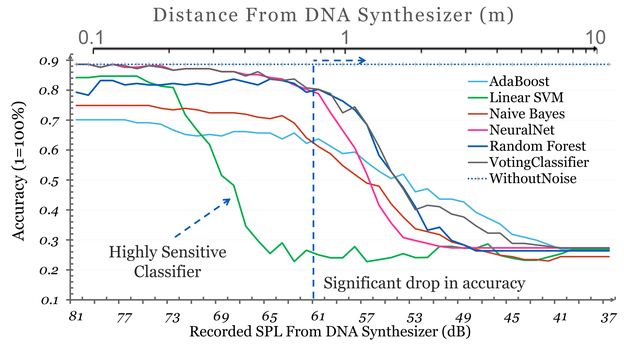
	Delivery #1	Delivery #2	Delivery #3
A	0.9	0.03	0.12
G	0.05	0.8	0.4
C	0.01	0.15	0.35
T	0.04	0.02	0.13

- **Problem Statement:** If a predicted sequence is determined to contain errors, what is the next best alternative?
- **Solution:** Find the next most probable sequence:
  1. Map the type prediction probabilities to a DAG.
  2. Find the K longest paths from the Start to End node.
- **Complexity:** O ($nlogn + k$) compared to O ($n4^n$)

(6) Results







Test Cases

Case #	Original Oligonucleotide sequence / Predicted Oligonucleotide sequence	Sequence Length	Accuracy (%)	BLAST match	Number of guesses to have N or less mispredicted amino acid N=3	N=2	N=1	N=0	Brute Force Complexity
1	CGCAA**G**TACTCCTG**C** CGCAA**T**TACTCCTG**A**	15	86.67	Yes	1	1	3	21	15x4^15
2	GGAATAGTAGAAG**AA**TGCTGCACAA**G**CATATGCAGCCTA**T**ACGAACTAGAAGAC**T**ACTGCGAC GGAATAGTAGAAG**CG**TGCTGCACAA**T**CATATGCAGCCTA**C**ACGAACTAGAAGAC**G**ACTGCGAG	63	**90.48**	Yes	12	29	>100	>100	63x4^63
3	TGGCGACAT**G**ATAACCCGTCGGA**G**GATCCGGG**G**CG**G**G**GG**CACCTC TGGCGACAT**T**ATAACCCGTCGGA**T**GATCCGGG**T**CG**T**G**TT**CACCTC	45	77.78	Yes	36	>100	>100	>100	45x4^45
4	TTTT**T**CGACCGGT**A**T**GA**T**T**CCGCCCGTGACCCAGGACGCTTGCTT TTTT**G**CGACCGGT**C**T**TC**T**G**CCGCCCGTGACCCAGGACGCTTGCTT	45	88.89	Yes	1	3	35	>100	45x4^45

- **88.07%** average random base classification accuracy.
- Robust to common noise in the environment.
- Microphone distance can be increased up to **0.7 meters**.
- Matching test cases results.
- Postprocessing tools such as **BLAST** compensate for errors in predictions.

(7) Broader Impact

- This work raises awareness among the bioengineering community to consider the possibility of a new set of attacks against the confidentiality of DNA synthesizers.
- Similar attack methods could potentially be used to breach the confidentiality of other information sensitive bioCPS tools.
- We show the potential for a government agency to non-intrusively monitor the synthesis of DNA by malicious parties to prevent large scale bioterrorist attacks.

(8) Related Publication

- Faezi, S. , Chhetri, S. R., Malawade, A. V., Chaput, J., Grover, W. H., Brisk, P., and Al Faruque, M. A.. “Oligo-Snoop: A Non-Invasive Side Channel Attack Against DNA Synthesis Machines” accepted to be published in The Network and Distributed System Security Symposium (NDSS’ 2019)