Responsible use of Artificial Intelligence to Enhance Society

Rajarshi Gupta

Head of Artificial Intelligence, Avast

Garry Kasparov

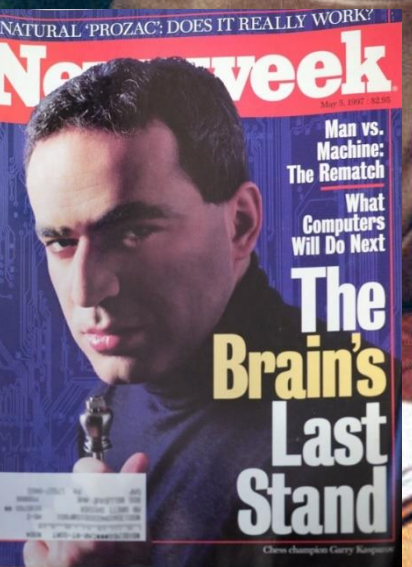
Chess Grandmaster &
Avast Security Ambassador

Garry Kasparov

Avast Security Ambassador

1985

1997


NATURAL 'PROZAC': DOES IT REALLY WORK?
Newsweek
Man vs. Machine: The Rematch
What Computers Will Do Next
The Brain's Last Stand

avast

the CONNECTED CAR

avast


Data Breach
Cyber Attack
DeepAttacks
Protection Failed
System Safety Compromised

avast

Rajarshi Gupta

Head of Artificial Intelligence, Avast

Data, data, data: What drives the world's largest consumer security network

MACHINE LEARNING

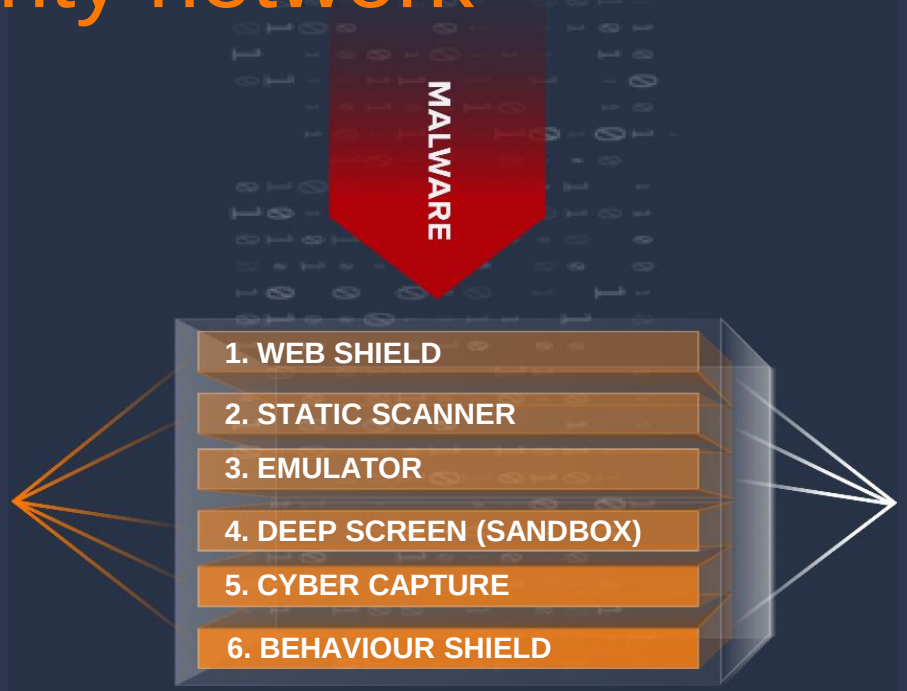
Every Month, Avast:

Handles 30+ million new executable files, 25 percent of which are usually malicious.
Continuously sifts through 390TB of quality security data

Monthly Engagement

290M+

145M+



AVAST CLOUD ENGINE

Every Month, Avast:

Prevents +2 billion malware attacks

Pushes 50 PB of data

Monthly Engagement

> 10,000 Servers

Across 10 Locations Worldwide, processing monthly:

> 300M Files

> 200Bn URLs

Dynamic malware detection engine

Purpose-built approach that takes < 12 hours to add new features, train, and deploy into production

COLLECTION

Goal: Harness as much data as possible

Avast Advantage: 6X more consumer PC users than the nearest competitor(1)

EXTRACTION

Deconstruct data into billions of artifacts

Proprietary Local Expert architecture leverages over **500+ features** (e.g. size, origin, age, and file entropy)

TRAINING

Update models to understand the intention of a sample

New models can be trained on the entire historical dataset in less than 12 hours

EXECUTION

Precisely and quickly identify what is benign vs. malicious

Endpoint-based models are updated 200+ times per day

AI in Action: Using Neural Nets to Optimize the Classification Engine

Published at ICLR2018

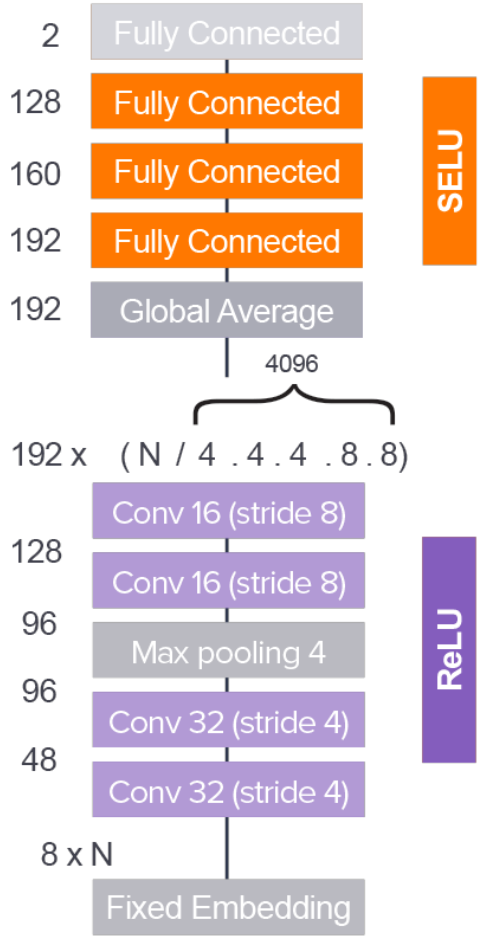
Goal: augment our traditional handcrafted models with machine-generated features

- Train a Convolutional Neural Net using the raw sequence of bytes from the binary files
- Training set of 20 million Windows PE files

Results

- Raw model achieves comparable accuracy to hand crafted features
- Choosing machine-generated features makes it much harder to evade
- Enriched features model shows extra gain of using both sets of features

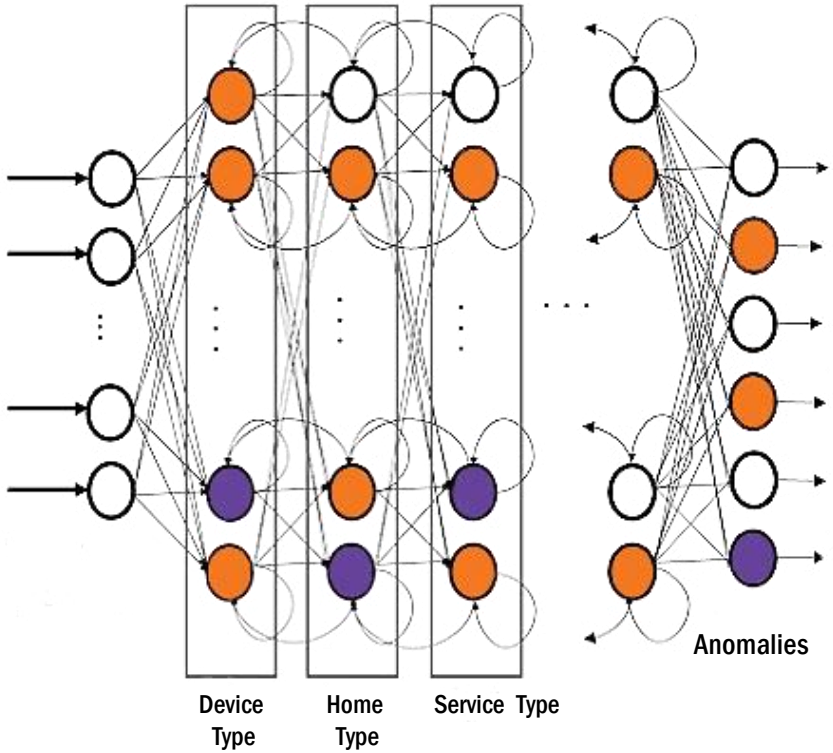
CLASSIFIER	ACCURACY
MalConv	94.6%
Avast Raw Files ConvNet	96.0%
FNN on features crafted manually	96.2%
FNN on enriched features	97.1%



AI in Action: Deep Neural Net for IoT Traffic

Input: Flow statistics from millions of homes

Input device traffic information for many devices, in many homes over a long time period



Benign

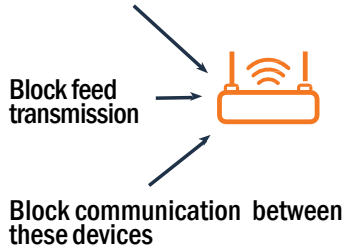
DDos: Many devices attacking same domain

Benign

Unexpected destination for baby monitorfeed

Benign

Unexpected traffic between devices within a home



Multi-Level Model

DEVICE TYPE

IoT devices have very limited behavior

Identifying devices allow us to model their behavior

SERVICE TYPE

Many devices plus internet makes up services, e.g. Netflix

ATTACK TYPE

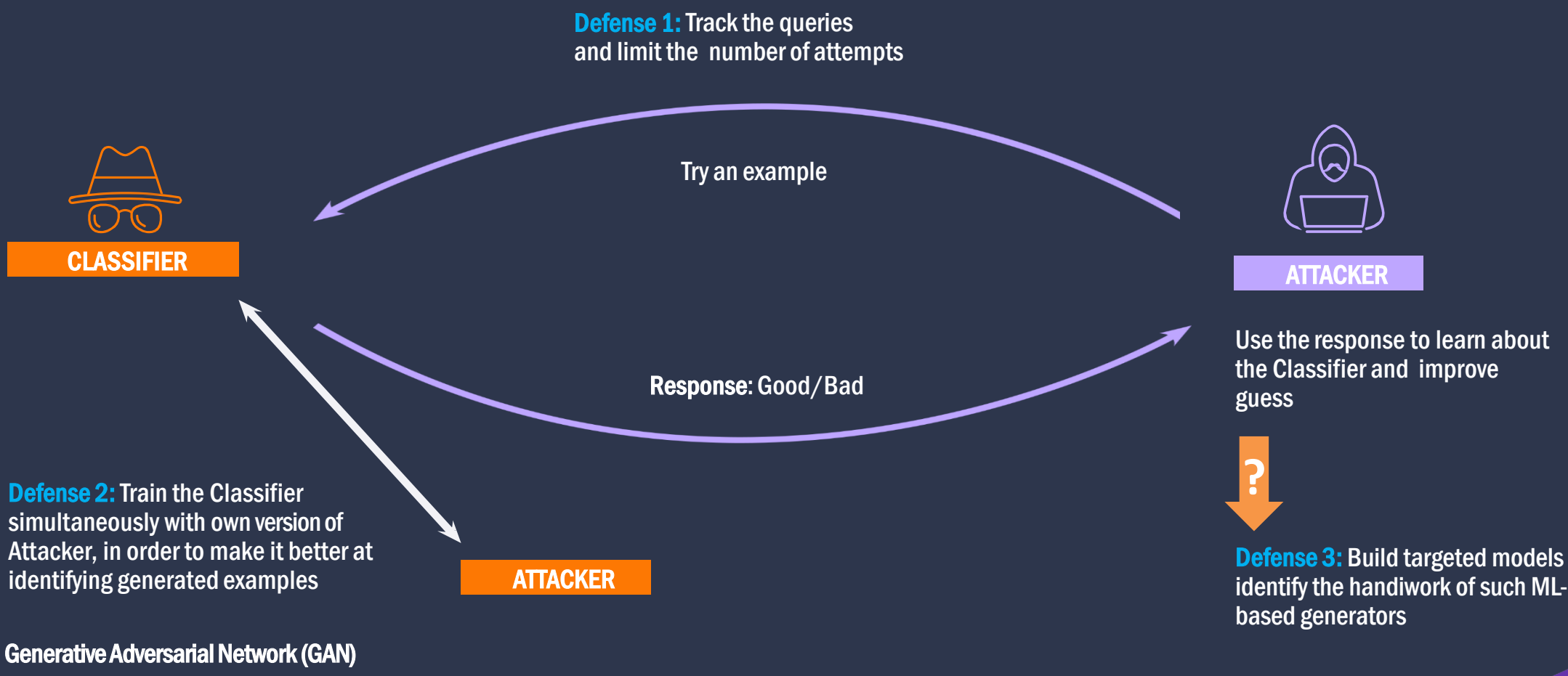
May focus on a device type, or servicetype

Output: Autonomously identify anomalous traffic

- Recognize unknown attacks
- Identify the device or service causing the attacks

avast

Using AI to Defend Against AI-based DeepAttacks



avast


avast
Thank You

Smart Home Networks in Mexico

In Dec 2018:

- Avast scanned 2.5M devices in 390K homes
- 47% of homes have at least one vulnerability
- 28% of homes have at least one router vulnerability
- 7% of home routers have weak WiFi password

AI is Security's Greatest Opportunity

VOLUME **SCALABILITY**

Coping with the sheer volume of new threats would be impossible without ML

VELOCITY **SPEED**

Most threats have very short longevity; machines can act much faster

VARIETY **ACCURACY**

ML is also really good at taking into account large amounts of contextual data

avast