When a Family of Iris Flower is Normal, Then are Others Abnormal?

Akira Imada

Brest State Technical University
Belarus

Network Intrusion Detection

When we design such a system,
by means of a machine learning technique
we need an *artificial* data to train and to test the system.

$$\Downarrow$$

Is this method reliable?

(Cont'd)

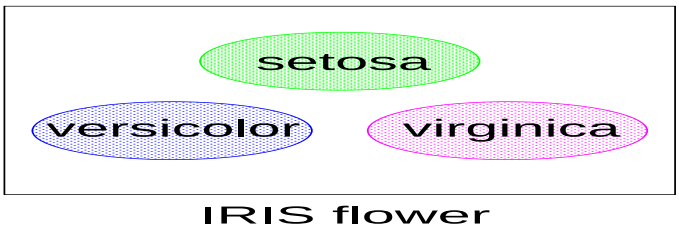
Specifically, is the frequently used *"Iris Flower Data"* useful for the purpose?

This talk is not a report of success,

but a challenge to those who have claimed success.

What is iris flower dataset?

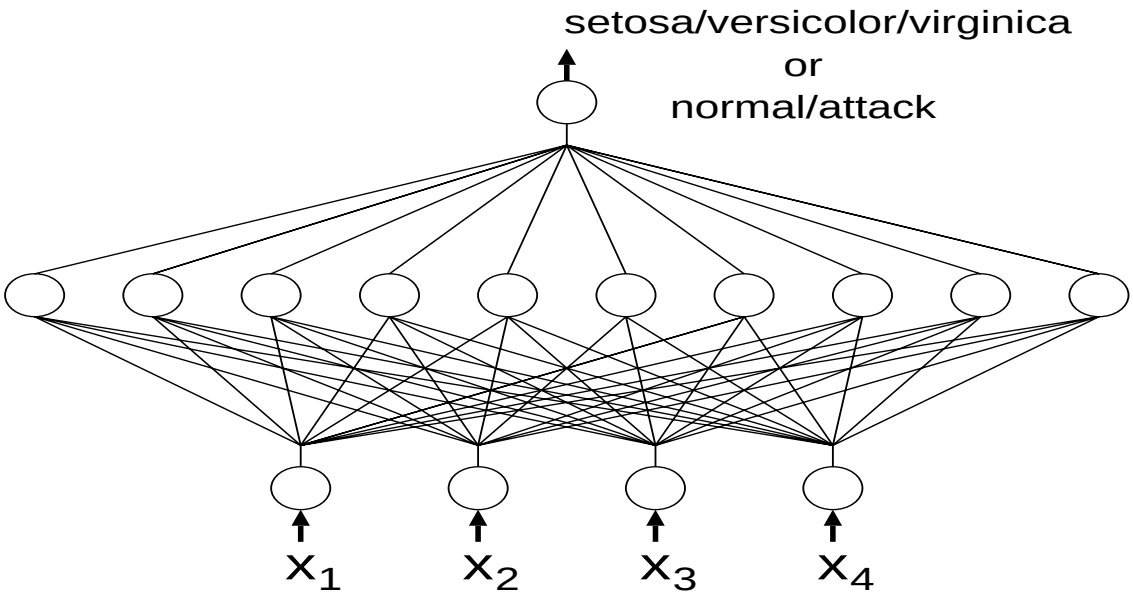
3 families of iris flower;
Each family has 50 samples $\Rightarrow$ 150 samples in total;



Each with 4 features (length & width of sepal & petal) like

(0.35 0.28 0.46 0.02 Setosa)

Typical implementation by a neural network



Training with GA as an example

chromosome:
$(w_1, w_2, w_3, \cdots\cdots, w_N)$

To evaluate fitness:
Firstly, we give a training set of normal and attack samples *to each of the individuals.*
Then ratio of correct outputs is the fitness value.

10-fold cross validation

The whole 150 samples are divided into 10 parts,

$\Downarrow$

so that each has 15 samples uniformly drawn from 3 classes.

$\Downarrow$

The system is *trained* by the remaining 135 samples.

$\Downarrow$

Then the 15 samples originally picked up are used to *test.*

$\Downarrow$

This is repeated 10 times.

2-fold cross validation for simplicity here

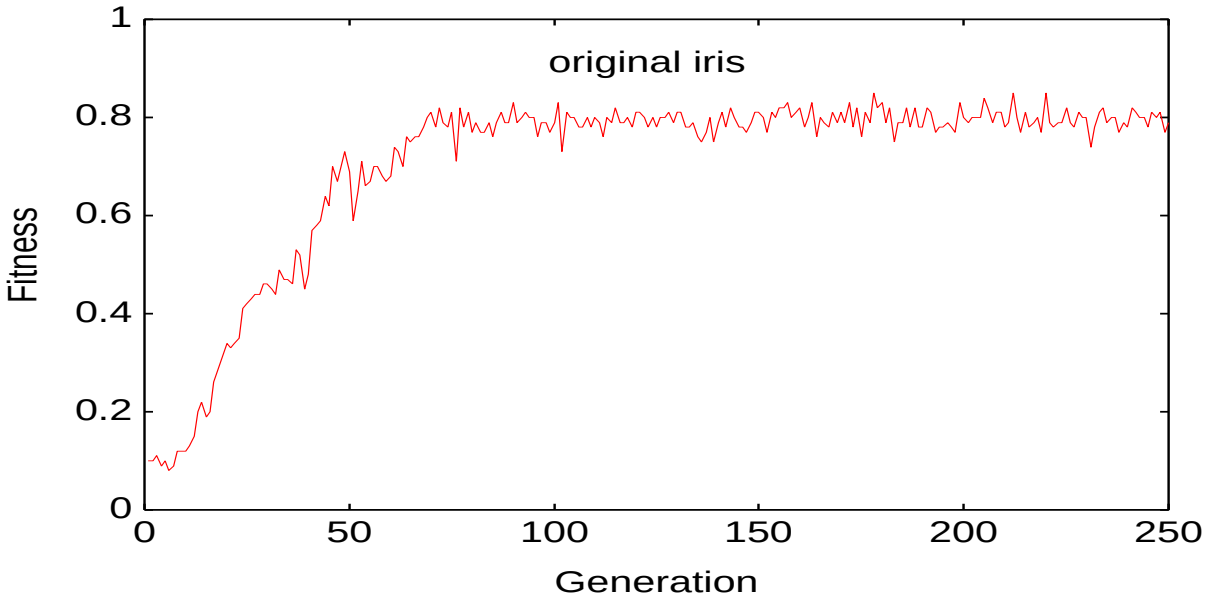
Train with a half of both of normal and attack.

$\Downarrow$

Test with the remaining half.

Our GA experiment – three families of iris

Evolution of average fitness



Reference

Iris-flower & Network Intrusion Detection

Kim, J. and Bentley, P., (1999a), " The Human Immune System and Network Intrusion Detection. "7th European Conference on Intelligent Techniques and Soft Computing (EUFIT '99), Aachen, Germany.

Kim, J. and Bentley, P., (1999b), " The Artificial Immune Model for Network Intrusion Detection, 7th European Conference on Intelligent Techniques and Soft Computing (EUFIT '99), Aachen, Germany.

Kim, J. and Bentley, P., (2000), " Negative Selection within an Artificial Immune System for Network Intrusion Detection", the 14th Annial Fall Symposium of the Korean Information Processing Society, Seoul, Korea.

Yu Guan et al. (2004) "K-means+: An Autonomous Clustering Algorithm" University of New Brunswick, Fuculty of Computer Science Technical Reports, TR04-164

G. G. Grinstein et al. (2002) "Benchmark Development for the Evaluation of Visualization for Data Mining" Information Visualization in Data Mining and Knowledge Discovery, Morgan Kaufmann, San Francisco, pp. 129-176.

D. E. Goodman et al. An Investigation into the Source of Power for AIRS, an Artificial Immune Classification System." Proceedings of International IEEE Joint Conference on Neural Networks, pages 1678-1683.

Emma Hart (2002) "Immunology as a Metaphor for Computational Information Processing: Fact or Fiction ?" Unpublished dissertation.

J. Timmis et al. (2004) "An Overview of Artificial Immune Systems." Computation in Cells and Tissues: Perspectives and Tools for Thought, Natural Computation Series, pp. 51-86. Springer, November 2004.

I.Antoniou et al. "Information processing in Artificial Immune System: Approaches and State of the Art."

Z. Ji and D. Dasgupata (2004) "Augmented Negative Selection Algorithm with Variable-Coverage Detectors." Proceedings of the Congress on Evolutionary Computation

G. Castellano and A. M. Fanelli(2000) "Fuzzy Inference and Rule Extraction using a Neural Network." Neural Network World Journal Vol. 3, pp. 361–371.

From success reports so far – 1

Castellano et al. (2000)
assumed one to be attack data while the other two normal.

(by Fuzzy-NN with T-S model on Iris-data)

$\Downarrow$

- successful attack detection rate = 96%
- false alarm rate = 0.6%

Yet another success report so far – 2

Kim & Bentley (2001) claimed

(by Artificial Immune Model on Iris-data)

$\Downarrow$

- Successful Detection Rate reached 100%
- False Alarm Rate was only 1%.

But can we be so optimistic?

$\Downarrow$

"When a family of iris is normal then are others abnormal?"

Let's try a thought-experiment an extreme case:

What if we design

always-answer-yes-machine

or

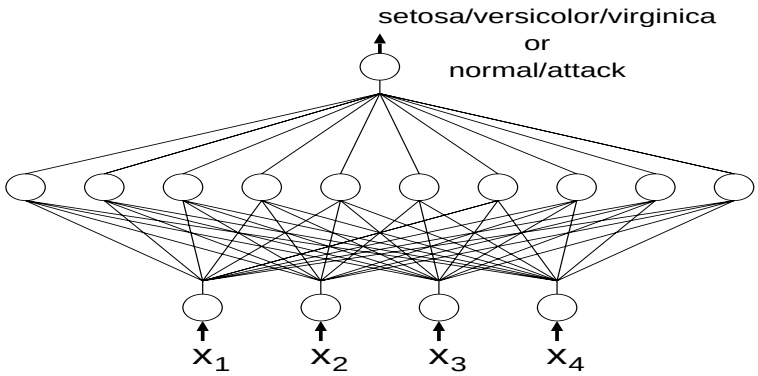
randomly-answer-machine?

$$\Downarrow$$

Detection rate of normal $\approx 100/150 = 66\%$.

(Cont'd)

Evolving NN by GA often try to create this *"always-output-yes-NN"*



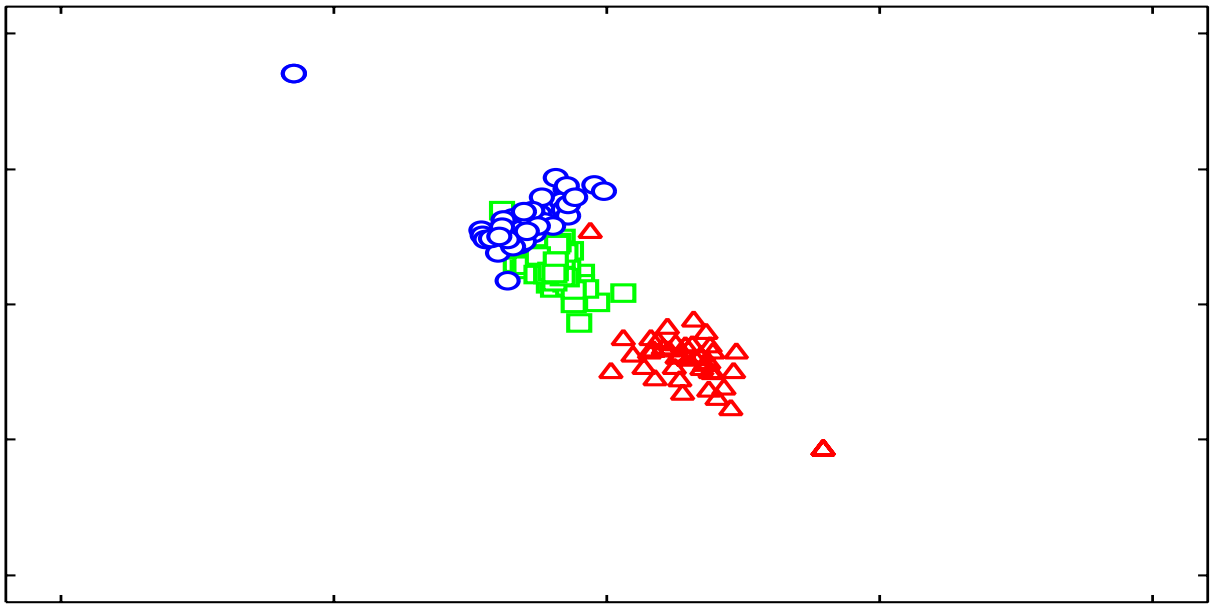
Let's visualize iris families — Sammon mapping

Sammon mapping maps points in a high-D space into 2-D

by

keeping distance relation preserved as much as possible
or
approximating distances in the high-D space
by distances in 2-D space with a minimal error.

A Sammon mapping of iris families



How is a hucker like? (From a newspaper)

"Those highly qualified hackers who provide security services to companies during the daytime

and then

go home at night to conduct totally illegal hacking

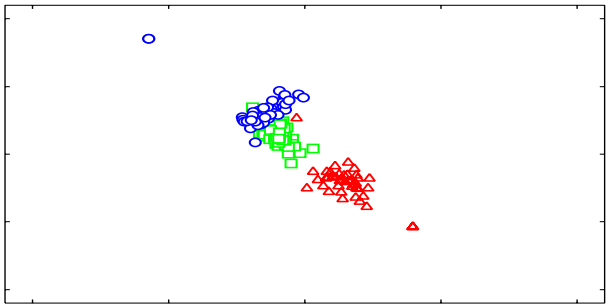
are

the ones who are the most dangerous."

– by Enis Senerdem from Turkish Daily News on 29 March 2006.

Where do outliers lie?

Not simply in the domain for the other families,
like IRIS data!!



⇓

Not at the point at random, either.

But outlier usually hides behind Normal.

Standard Experiment on Iris-data

Train and test with
one family as attack while the other two as normal

Train with a part of the data
$\Downarrow$
Test with the remaining data

Outliers lie:

Not in the domain for the other families

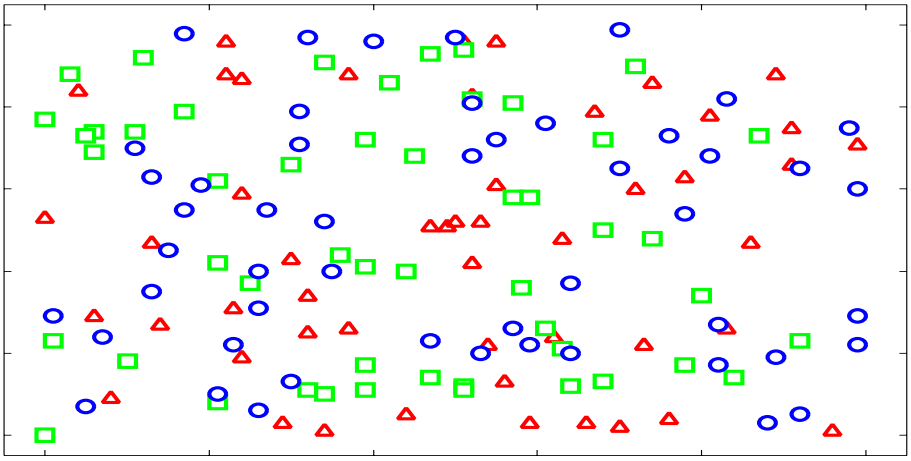
⇓

Not at the point at random, either.

⇓

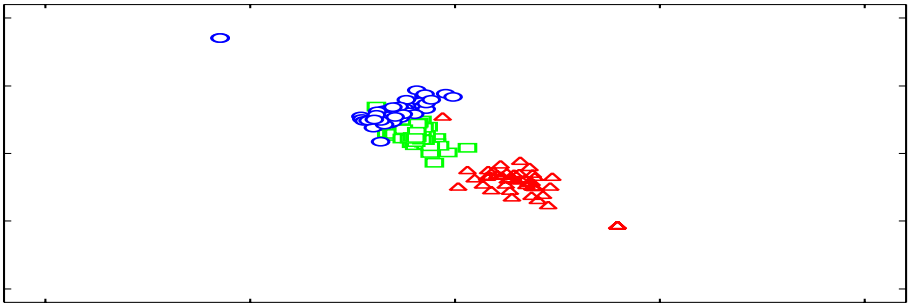
But outlier usually hides behind Normal.

What if the 3 families were distributed like below?



(Cont'd)

This doesn't look like a reality, but locally, normal and attack data are distributed in this way, more or less.



Challenge I

Randomly located normals and attacks

(assuming both normal and attack don't create cluster)

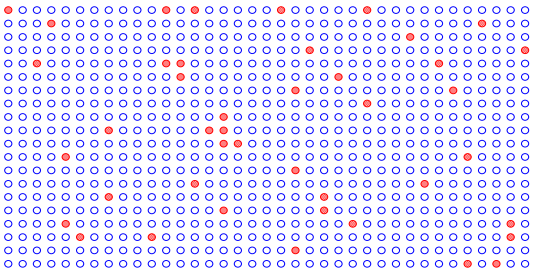
Create 100 normal samples and 50 attacks
all at random.

Train with half of normal and attacks.
$\Downarrow$
Test with remaining half.

Ayara et al. (2002)

"Negative Selection: How to Generate Detectors."

Random attacks in 8-bit binary universe



152 random attacks out of $2^8 = 256$ search points.

$\Downarrow$

Asserted that successfully trained and detected them.

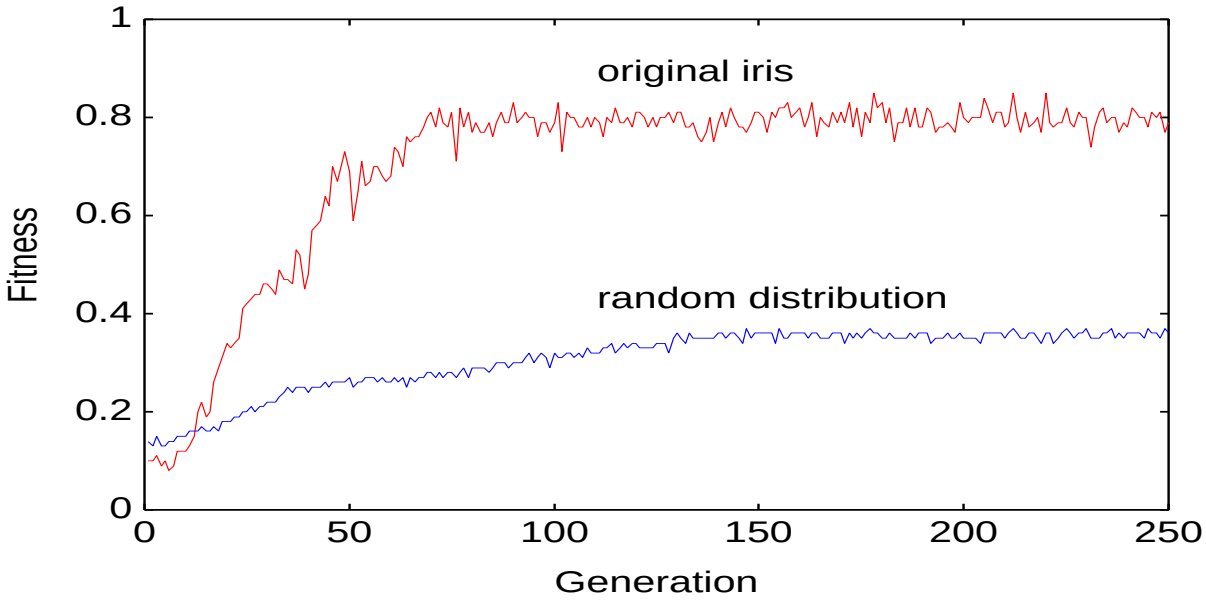
(Cont'd)

This is not so difficult due to small universe.

On iris dataset,
this is extremely difficult, if not impossible.

Our GA experiment – random normal & attack

Evolution of average fitness



Outliers lie:

Not in the domain for the other families.

$\Downarrow$

Not at the point at random, either.

$\Downarrow$

But outlier usually hides behind normal.

Challenge II

What if attacks hide behind normals?

Create 100 normal samples again all at random.
Then 50 out of those samples are mutated as attack.

Train with half of the normal and mutants behind them.
$\Downarrow$
Test with the remaining half.

(Cont'd)

This is even more difficult than the previous one.

(Cont'd)

Let's now go back to Iris Data again.

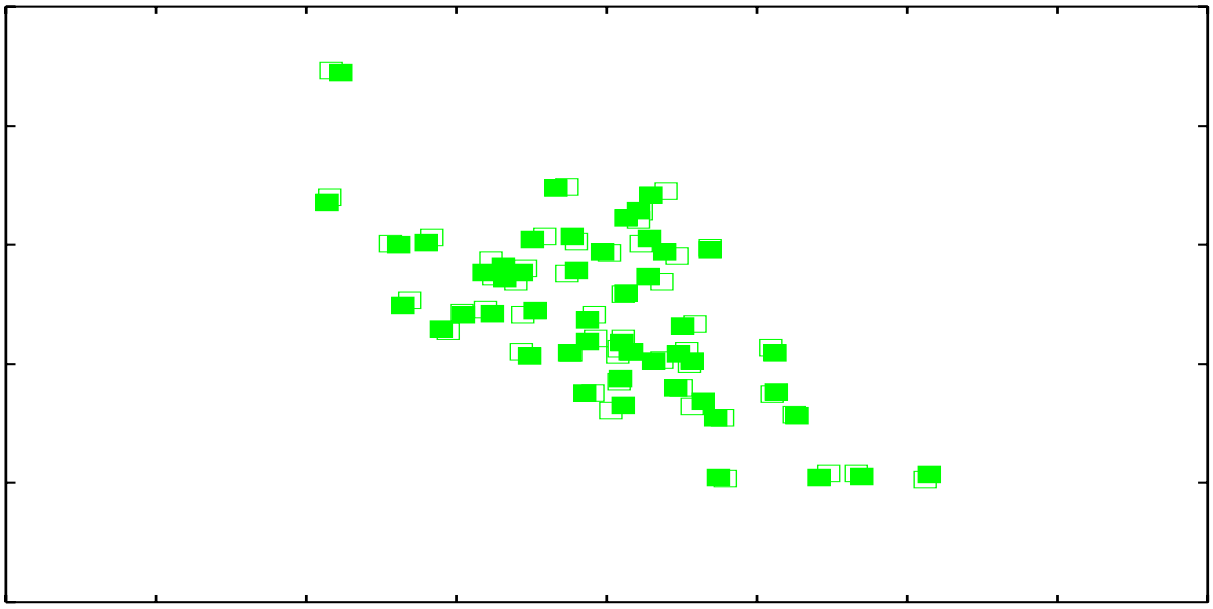
Challenge III

Attacks who hide behind normal iris family

Assume one family as normal.
Then those samples are mutated as attack.

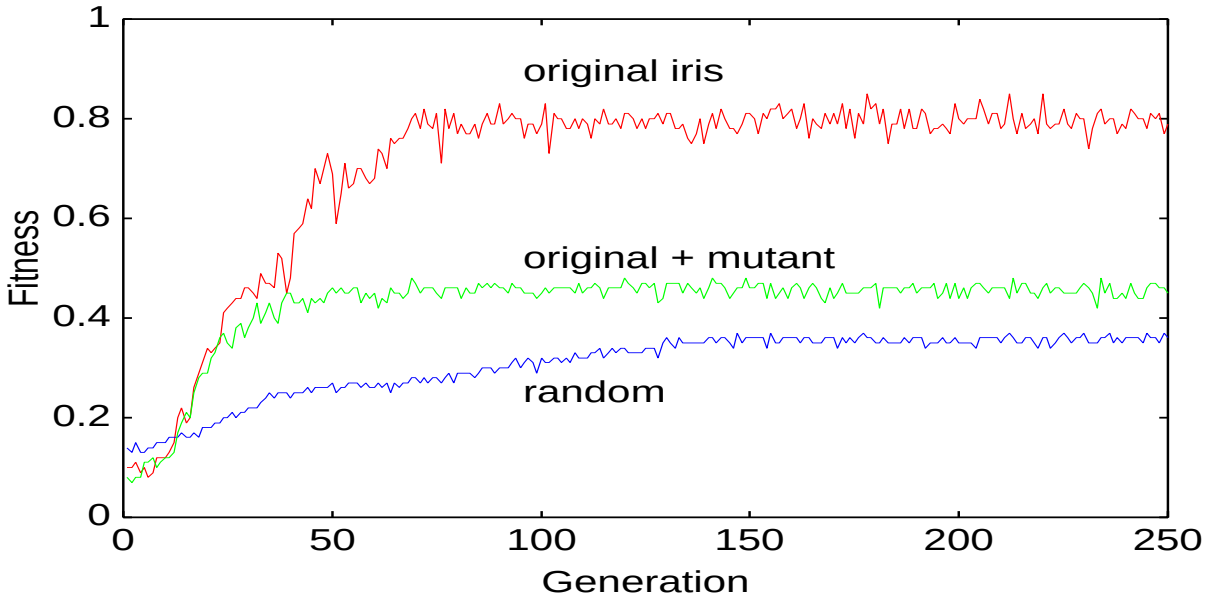
Train with half of normal and mutated attacks.
$\Downarrow$
Test with the remaining half.

One family and its mutants



Our GA experiment – normal & mutant

Evolution of average fitness



Challenge IV

Can we train only with normal?

(This is actually is our MUST)

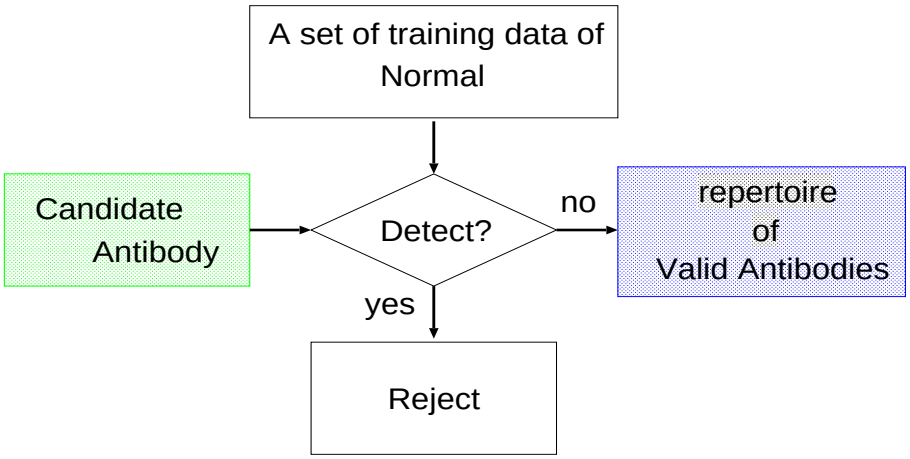
Again one family is attack while the other two normal.

Train *ONLY* with normal.

$\Downarrow$

- Test-1: with the other family as attack.
- Test-2: with the randomly specified data as attack.

An immune approach – Negative Selection



Exhaustive detector ... Forrest, Perelson et al. (1994)

A bottleneck of the negative-selection algorithm

Size of candidate antibodies
increases exponentially
with size of normal.
(Ayara et al.)

From success reports so far – 3

Gomez et al. (2003)

"A set of fuzzy rules characterized abnormal space using only normal samples."

(with 10% dataset of KDD-cup-99).

⇓

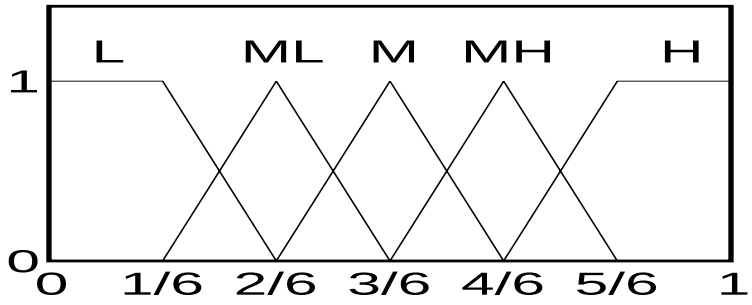
"It detects attacks with the detection rate 98.30% and false alarm rate 2.0%."

Let's apply Gomez' idea to iris data

The 5 fuzzy linguistic terms

$\Downarrow$

$\{$*Low, Medium-low, Medium, Medium-high, High*$\}$



(Cont'd)

An example of a set of rules

IF
x_1 *is Law or Medium-law;*
x_2 *is High;*
x_3 *is Medium or Medium-high or High; and*
x_4 *is Medium-low or Medium.*
$\Downarrow$
THEN $\mathbf{x}$ *is ATTACK*

(Cont'd)

Assume we have k such rules.
And then,
a degree of membership of $\mathbf{x}$ *to the attack*
is calculated for given $\mathbf{x}$.

(Cont'd)

Gomez's chromosome in the context of iris data

E.g., ((11001)(01001)(10000)(11100)) means

$$x_1 \textit{ is } L \textit{ or } ML \textit{ or } H$$
$$x_2 \textit{ is } ML \textit{ or } H$$
$$x_3 \textit{ is } L$$
$$x_4 \textit{ is } L \textit{ or } ML \textit{ or } M$$

Fitness is evaluated as
a degree of membership of $\mathbf{x}$ *to the attack*
by giving N normal samples

(Cont'd)

Does this still work
on
random or *mutated* normal & attack data?

Challenge V

A metaphor of placebo — dummy pill

(necessary experiment by pharmaceutical companies)

"Drugs to ... are

no more effective than placebos for most patients."

found no significant difference than the placebos."

or

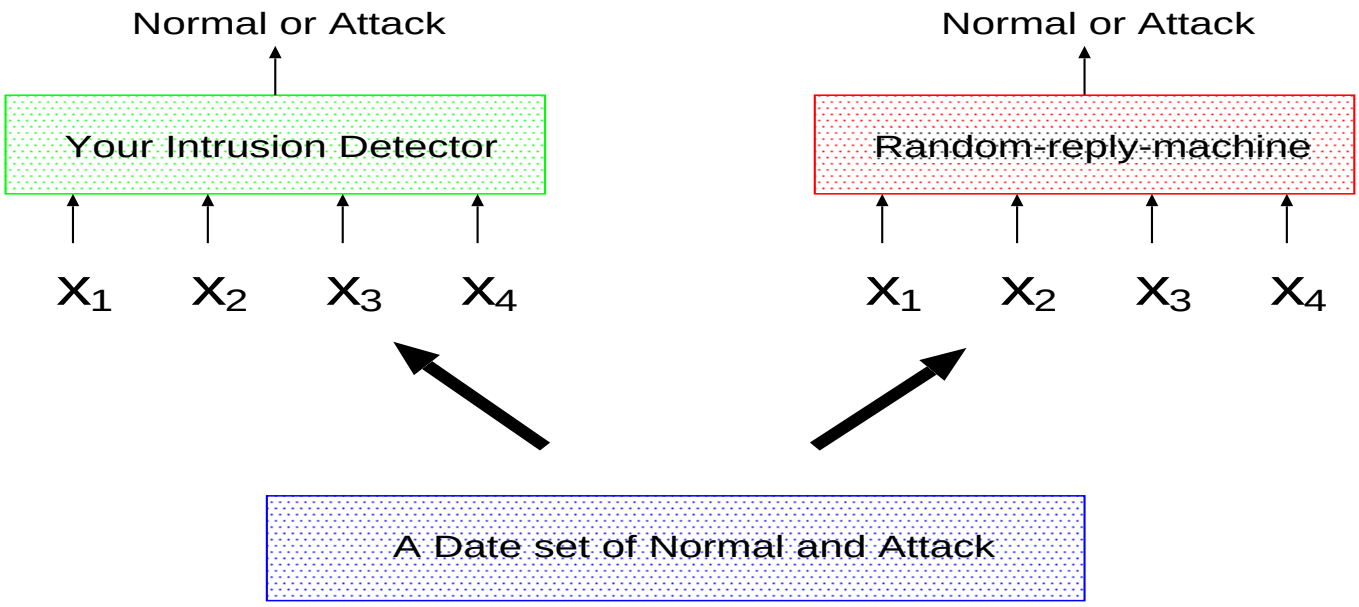
at increased risk of death compared to placebo."

– by Benedict Carey from New York Times on 12 October 2006.

(Cont'd)

We sometimes have a bias toward a too optimistic apriori expectation.

(as a metaphor of dummy pill)
Create a simple device
which randomly returns either normal or attack
regardless of the input.


Normal or Attack
Your Intrusion Detector
X1
X2
X3
X4
Normal or Attack
Random-reply-machine
X1
X2
X3
X4
A Date set of Normal and Attack

From success reports so far – 4

Joshi et al. (2005)

(by Hidden Markov Models on KDD-cup-99)

⇓

- 79% accuracy in correctly detecting attacks,
  - 21% is accounted for false positive rate

What if data comprises 80% normal and 20% attack?
(This is of usual situation in KDD-cup-99 dataset)

Concluding Remarks

- Can we successfully train the system with randomly distributed attack and normal?
- Can the system recognize a mutant of normal as an attack?
  - · with randomly located normal.
  - · with one family of iris.
- Can the system trained only by normal recognize attack?
- Is the performance better than a *random-replying-machine?*

(Cont'd)

We are trying all of the five challenges with many methods:
NN, GA, AIS, C4.5, Fuzzy-classifier ... etc.

But so far not so successful
except for the original experiment.

Needless to say,
no any intention of discouraging but a challenge.
Positive attack to this *pessimistic* talk will be welcome.
And of course a cooperation too.
We have many data-set immune to be successfully tested.