When an iris flower is normal then are others?

Akira Imada
Brest State Technical University, Belarus
Moskowskaja 267, 224017 Brest, Belarus
akira@bstu.by

Leonard da'Vinci
Graduate School of Information Science
Nara Institute of Science and Technology
8916-5 Takayama, Ikoma, Nara,
630-01 Japan
araki@is.aist-nara.ac.jp

***Abstract*— In the context of Network Intrusion Detection, we test a lately reported technique which generates a set of fuzzy rules to recognize unknown abnormal patterns using a test-function, what we call a-tiny-island-in-a-huge-lake. Our concern is whether or not we can train the system only with a set of already known normal patterns. Yet another of our concern is what happens in an extreme case where a sample of abnormal patterns are extremely few comparing to the normal ones, and what if it eventually shrinks to zero, which is what they call a-needle-in-a-haystack.**

I. Introduction

This paper reports a snapshot of our on-going experiments in which a common target we call *a-tiny-island-in-a-huge-lake* is explored with different methods ranging from a data-mining technique to an artificial immune system. Our implicit interest is a network intrusion detection, and we assume data floating in the *huge lake* are normal while ones found on the *tiny island* are abnormal.

A. How to put figures?

Our goal here is twofold. One is to know *whether or not it is possible to train a system using just normal data alone.* The other is to study *a limit of the size of the detectable area*, when we decrease the size of the island eventually shrinking to zero, equivalently so-called *a-needle-in-a-haystack* (See Fig. 1) which is still an open and worth while tackling problem.

B. This is an example

To learn these two issues, a fuzzy rule extraction system with fixed triangle/trapezoid membership functions are exploited in this paper.

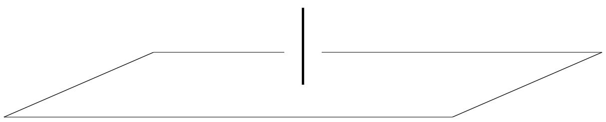

Fig. 1. A fictitious sketch of fitness landscape of *a-needle-in-a-haystack*. The haystack here is drawn as a two-dimensional flat plane of fitness zero.

C. An example of table

When we think of a network intrusion detection, we have a large collection of normal patterns while the number of possible anomaly patterns we know is extremely few, which is of usual cases. See Table 1.

Table 1. Generation number at which the elitest individual in the generation attains fitness of 1.000, for 12, 13, 14, 15 and 16 patterns.

patterns	12	13	14	15	16
generation	354	1268	6003	6688	10795

D. How to cite a reference

Furthermore, we usually don't know what do anomaly patterns look like in advance [1]. It is usually too late when we know it. Hence, our second concern is whether or not we can train the system with only a set of normal patterns.

II. An example of equations

A set of fuzzy rules is used to cover the non-self patterns. As already mentioned, self/non-self cells are represented by n-dimensional real valued vectors each of whose coordinate lies in $[-1, 1]$.

$$\text{fitness}(R) = 1 - \max_{\mathbf{x} \in Self} \{ \min_{i=1,\cdots,n} \{ \mu_{T_i}(x_i) \} \}$$

which implies how the rule covers the non-self space.

III. CONCLUSION

We have described how we would be able to find a small island in a huge lake with a system training only using data in a lake. This is a metaphor in which a very few of unpredictable abnormal transaction patterns hidden in an enormous amount of normal patterns, in the context of network intrusion detection.

References

[1] A. Imada (2004) "How a Peak on a Completely-flatland-elsewhere can be Searched for?" Proceedings of Advanced Computer Systems (ACS) and Computer Information Systems and Industrial Management Applications (CISIM), Vol.2, pp. 171–150.