

# Intrusion Detection with Genetic Algorithms and Fuzzy Logic

Emma Ireland

Division of Science and Mathematics  
University of Minnesota, Morris  
Morris, Minnesota, USA

December 2013  
UMM CSci Senior Seminar Conference

# The big picture



# Outline

- 1 Background
- 2 Genetic Algorithm Implementation
- 3 Fuzzy Genetic Algorithm Implementation
- 4 Conclusions

# Outline

- 1 Background
  - Types of Networking Attacks
  - Detection Methodologies
  - Data Sets - KDD99 and RLD09
  - Rules
  - Fuzzy Logic
  - Genetic Algorithm
  - Determining the accuracy of an algorithm

2 Genetic Algorithm Implementation

3 Fuzzy Genetic Algorithm Implementation

4 Conclusions

# title here

# title

# KDD99

- Generated by simulating a military network environment in 1999.
- Has long been a standard data set for intrusion detection.
- Data in the set is classified as normal or attack activity.
- KDD99 uses 41 features.
  - *Features* are properties of a *record*, (either an attack or normal activity), that are used to describe the activity.

# Some features of KDD99

- ➊ duration: length of the normal or attack activity in seconds.
- ➋ src\_bytes: number of bytes sent from source to destination.
- ➌ num\_failed\_logins: number of failed login attempts.
- ➍ root\_shell: returns 1 if root shell is obtained, else returns 0.
- ➎ num\_access\_files: number of operations on access control files.
- ➏ srv\_count: number of connections to the same service as the current connection in the past two seconds.
- ➐ serror\_rate: percentage of connections that have "SYN" errors.
- ➑ same\_srv\_rate: percentage of connections to the same service.



# RLD09

# title

# title

# title

# title

# Outline

- 1 Background
- 2 Genetic Algorithm Implementation
  - Algorithm Overview
  - Experimental Design and Results
- 3 Fuzzy Genetic Algorithm Implementation
- 4 Conclusions

# title here

# title

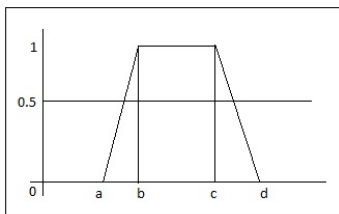


# Outline

- 1 Background
- 2 Genetic Algorithm Implementation
- 3 Fuzzy Genetic Algorithm Implementation
  - Fuzzy Algorithm
  - Algorithm Overview
  - Experimental Design and Results
- 4 Conclusions

# Measuring the probability of a record being an attack

- Trapezoidal shape



- The parameters are the values of a feature.

- Fuzzy algorithm

**if** data value is between  $b$  and  $c$  **then**  
     $\text{prob} = 1.0$

**else if** data value is between  $a$  and  $b$   
**then**

$$\text{prob} = (\text{data} - a) / (b - a)$$

**else if** data value is between  $c$  and  $d$   
**then**

$$\text{prob} = (d - \text{data}) / (d - c)$$

**else**

$$\text{prob} = 0.0$$

**end if**

# Encoding of features and rules

- The four parameters are encoded into blocks.
- Each block is a feature with values between 0.0 and 7.0.

010	011	100	101
a=2	b=3	c=4	d=5

- A rule has 12 blocks of features, at the end is the type of attack.

010	011	100	101	.....	010	011	101	111	DoS
a=2	b=3	c=4	d=5	.....	a=2	b=3	c=5	d=7	
Block 1					Block 12				Type

# title

# title

# Outline

- 1 Background
- 2 Genetic Algorithm Implementation
- 3 Fuzzy Genetic Algorithm Implementation
- 4 Conclusions**

# Conclusions



# Thanks!

Thank you for your time and attention!

## Questions?



# References