

Dempster-Shafer for Classification and Anomaly Detection

Qi Chen

Supervisor: Uwe Aickelin

2

Outline

- Introduction
 - D-S Classification and Anomaly Detection Applications in Literature
 - Our Approach
 - Aims and Scope
- The Dempster-Shafer (D-S) Theory
- D-S Based Classification Systems
- D-S Based Anomaly Detection Systems
- Conclusions

D-S Applications in Literature for Classifications and Anomaly Detection

- A Tool for Combination
 - At Decision Levels-combine results from various classification algorithms
 - each feature/resource use a classification algorithm as mass function, combine the results of various features/resources

3

Our Approach

- D-S is used as a framework to design classification systems
- No other classification algorithms are used

4

Aims and Scope

- How to apply D-S to real-world problems
 - What guideline should we follow
 - How to develop mass functions
 - How to select features
- IS D-S suitable for Classification
 - For Numerical Data
 - For Nominal Data
- Is D-S suitable for Anomaly Detection

5

Loose Guideline to Apply D-S

1. Define The Frame of Discernment(Θ)
2. Select Meaningful Attributes
3. Design Mass functions for each attribute
4. Design a DRC combination strategy
5. Select/Develop A Decision Rule

7

The Dempster-Shafer (D-S) Theory What is D-S

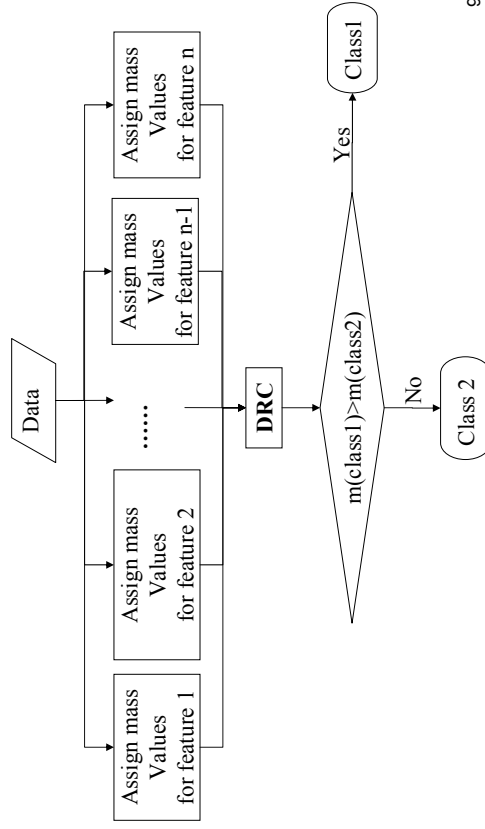
- A Mathematic Theory of Evidence
- Relevant Parts
 - Mass Function
 - Dempster's Rule of Combination (DRC)

6

- Classification Systems
 - Wisconsin Breast Cancer Dataset (WBCD)
 - Iris Plant Dataset
 - Duke Outage Dataset(DOD)- **Nominal Data**
- Anomaly Detection Systems
 - An Email Worm Detection
 - Ftp Server Exploit Detection (Wu_ftpd)

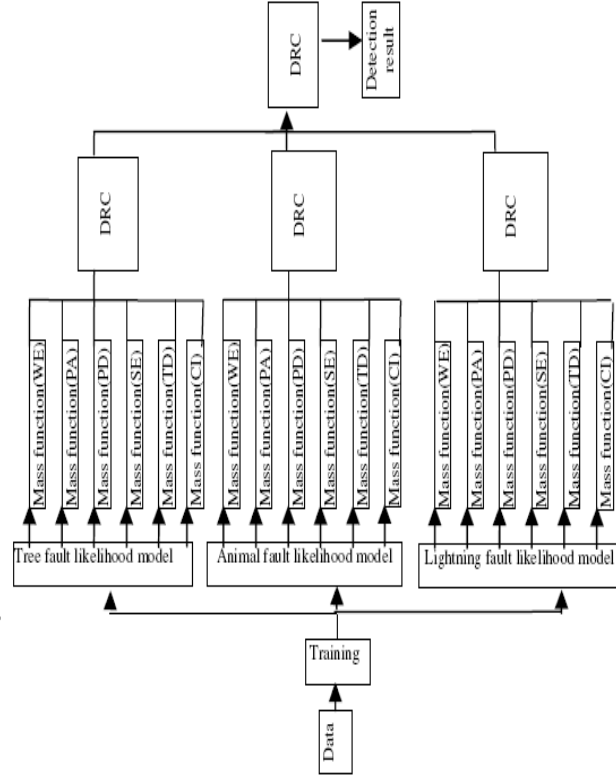
8

D-S System Architecture for One Step Classification Process



9

D-S System Architecture for DOD

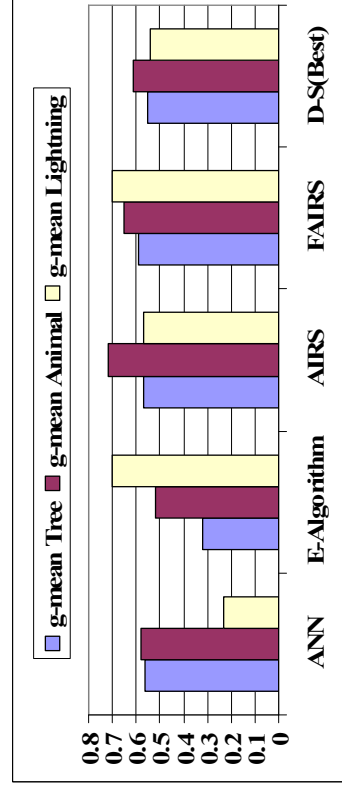


Results for WBCD, Iris Plant and Email

- Results of WBCD:
 - Classification rate = 97.6%
- Results of Iris Plant:
 - Classification rate = 95.47%
- Results of Email:
 - True Positive Rate=1, False Positive Rate=0.

10

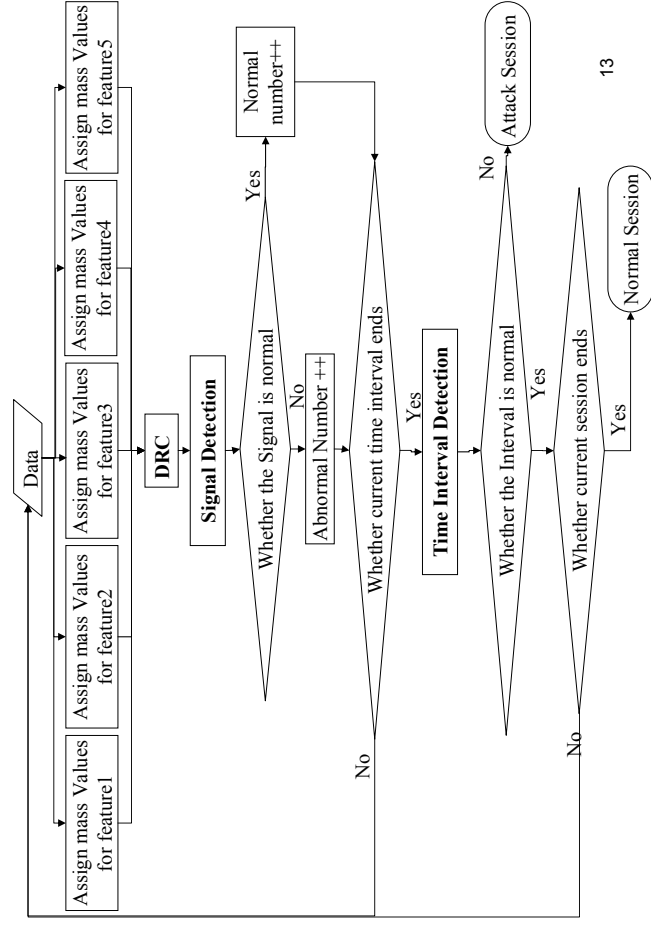
Results Comparisons for DOD



$$g - mean = \sqrt{TruePositiveRate \times TrueNegativeRate}$$

12

Architecture for ftp Exploit Detection



13

Conclusions-Part1

- Guideline for D-S Classification and Anomaly Detection Applications
 1. Define The Frame of Discernment(Θ)
 2. Select Meaningful Attributes
 3. Design Mass functions for each attribute
 4. Design a DRC combination strategy
 5. Select/Develop A Decision Rule
- Feature Selection

15

Results of Anomaly Detection Systems for Ftp Exploit Detection

	DCA	TLR	D-S
True Positives Rates	1.0	0.75	1
False Positives Rates	0.83	0.15	0.036

14

Conclusions-Part2

- Mass Function Development
 - Mass functions are derived from training data
 - mass functions based on likelihood measurements from the training data
- Mass Function Optimization

16

Limitations and Future Work

- For All Applications: Decision Rule
 - The belief values for singleton hypotheses are compared, the range [belief value, plausibility value] can be used
- For Duke Outage Application
 - Each model is viewed as the same with same trustiness, further work can be done to assign various trustiness to each model
- For Ftp Exploit Detection
 - System Call Information Can Be Used in Future Work

17

Thanks!

Question?

18