Traffic Analyzer

Advanced Multimedia Processing Lab

AMP

Prepared by: Devi Parikh

CMU Team:
Devi Parikh, Li-Hao Liu, Tsuhan Chen

Date: November 10, 2005

Carnegie Mellon

Overview

- Motivation
- Objective
- Algorithm
- System Layout
- Demonstration
- Conclusions
- Questions

Carnegie Mellon

Motivation

- Intrusion detection in computer networks
- Trade off between false alarm rates and ability to detect new attacks
- Pattern recognition can strike a balance

Objective

- Neural network based classification system
- Ensemble of classifiers approach
- Combine information from multiple sources (use most informative sources for different attack types)
- Adapt to changing traffic patterns

Algorithm

- Classification system
  - Subsections corresponding to each of the categories (normal, different attack types)
    - Subsections corresponding to each of the sources of information
      - Building blocks – Neural Networks

Algorithm

Input: Data split into **Training and Validation**

Output: Trained classification system

For each class

For each feature set (source of information)

Loop1: Train 1st NN on 2/3 of training data

Validation performance = V1

If V1<0.5 - Goto Loop1

Calculate weight

Loop2: Train 2nd NN on 2/3 of training data

Validation performance = V2

If V2<0.5 - Goto Loop2

Calculate weight

Validation of two networks combined = V12

If V12<V1 - Goto Loop2

End

End

Algorithm

Input: **Testing** dataset

Output: Classification decision

For each instance

Combine NN within sources – Weighted Sum Rule

Combine decisions within categories – Product Rule

If exactly one decision is positive

Pick corresponding class as decision

Elseif more than one are positive

Pick class with highest confidence

Elseif all decisions are negative

Raise an alert

Expert involvement to provide label

Update system by adding subsystem

End

End

Motivation

Objective

Algorithm

System - Layout

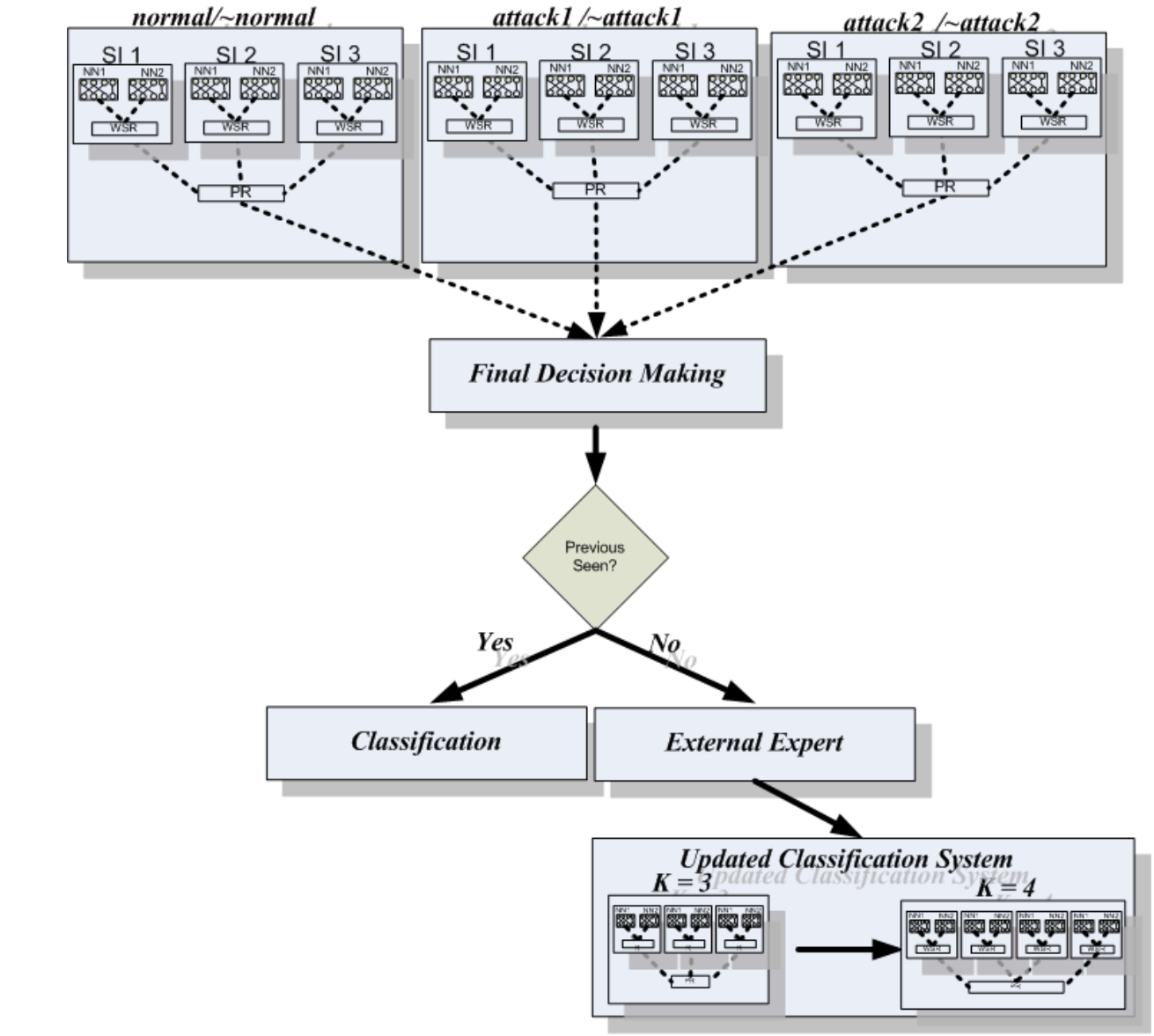
Demo

Conclusion

Questions

Carnegie Mellon

System Layout



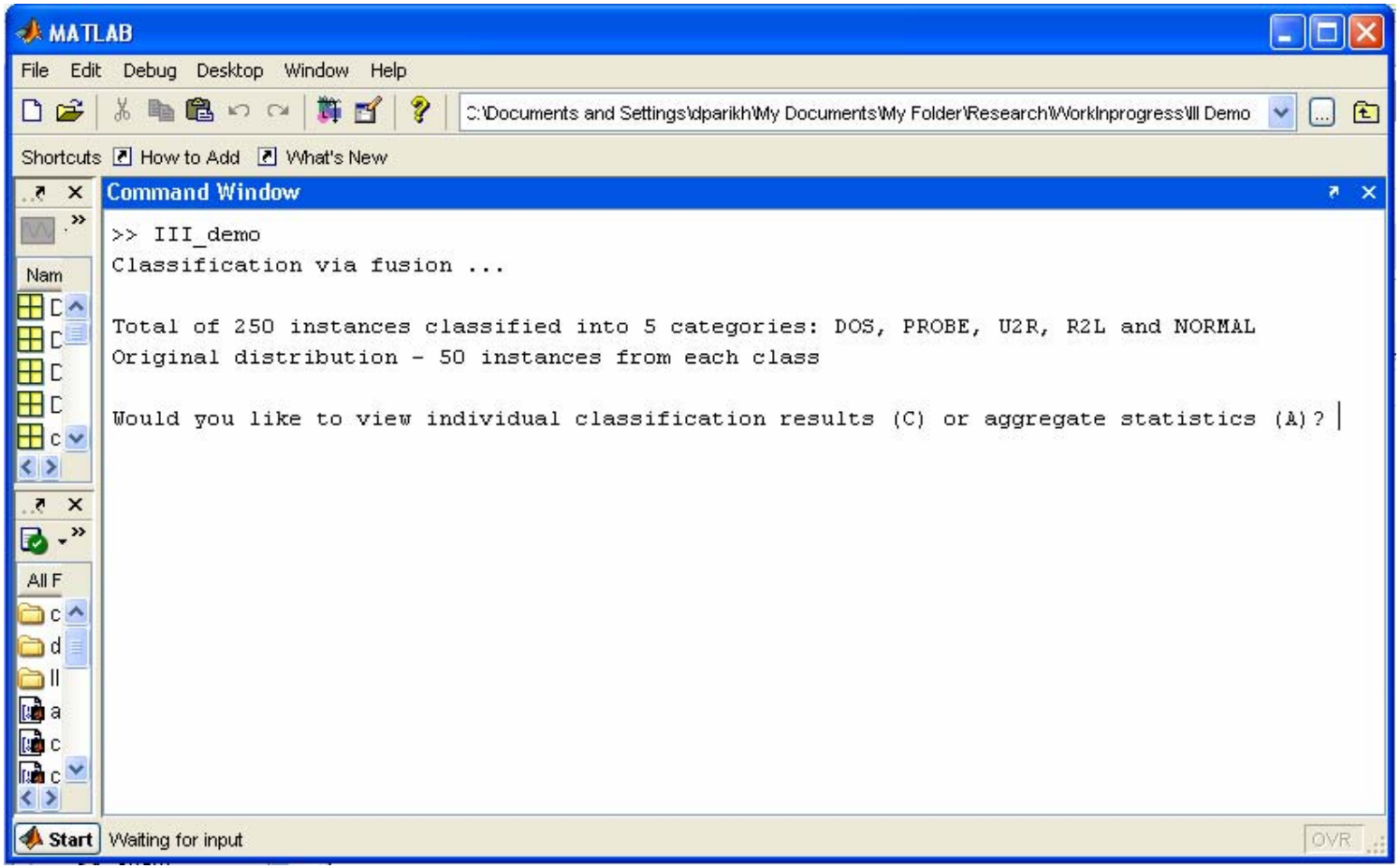
Demonstration

Demonstration (Testing) set up:

- KDD (MIT – DARPA, 1998) database
- 5 classes
  - Attack type 1: DOS
  - Attack type 2: PROBE
  - Attack type 3: U2R
  - Attack type 4: R2L
  - No attack: Normal
- Number of instances: 250 (50 per class)
- 3 sources of information
  - Source 1: Basic features
  - Source 2: Content features
  - Source 3: Traffic features
- Total of 30 neural networks in the classification system

- Motivation
- Objective
- Algorithm
- System - Layout
- Demo
- Conclusion
- Questions

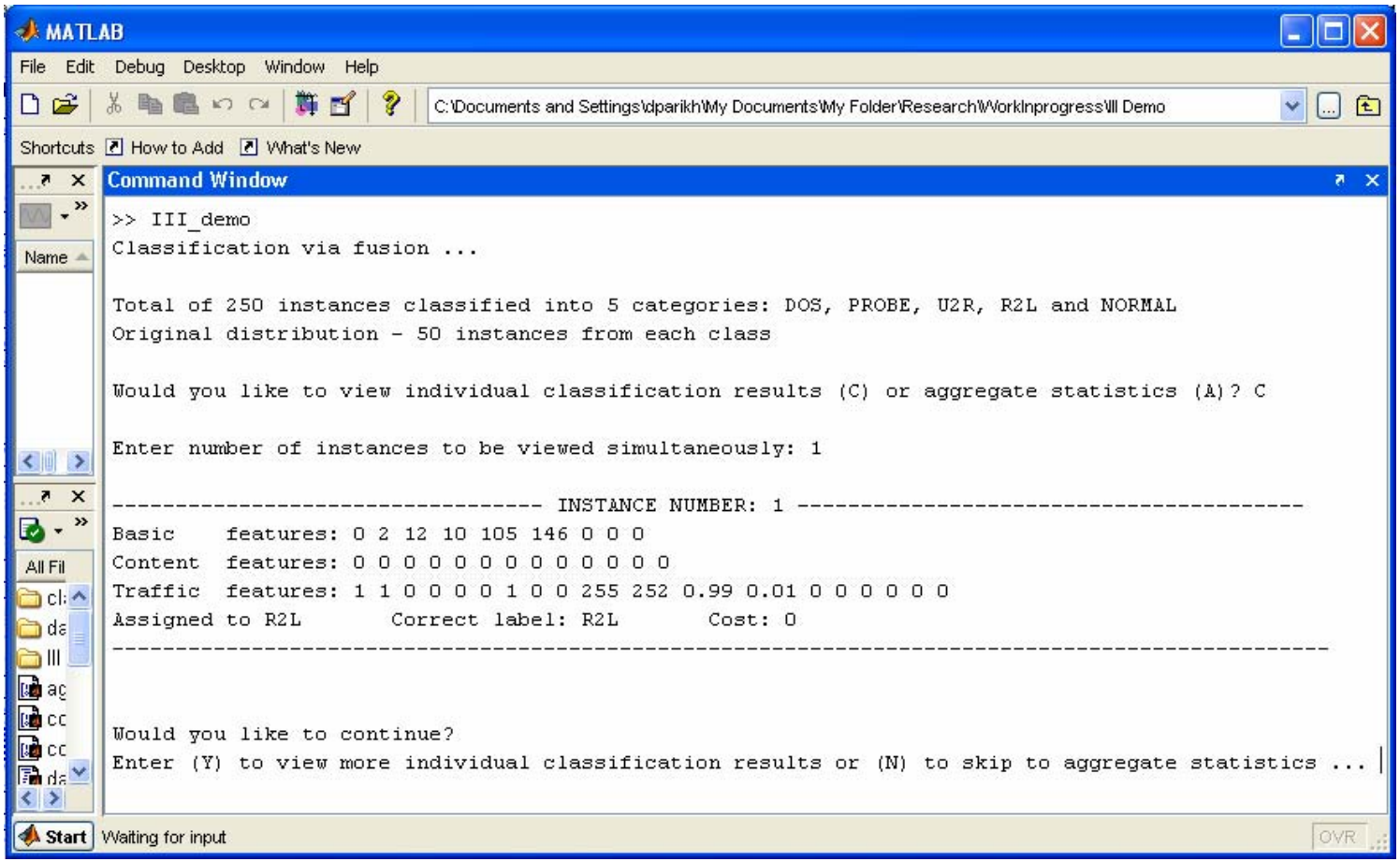
Demonstration



Carnegie Mellon

- Motivation
- Objective
- Algorithm
- System - Layout
- Demo
- Conclusion
- Questions

Demonstration



Carnegie Mellon

Motivation

Objective

Algorithm

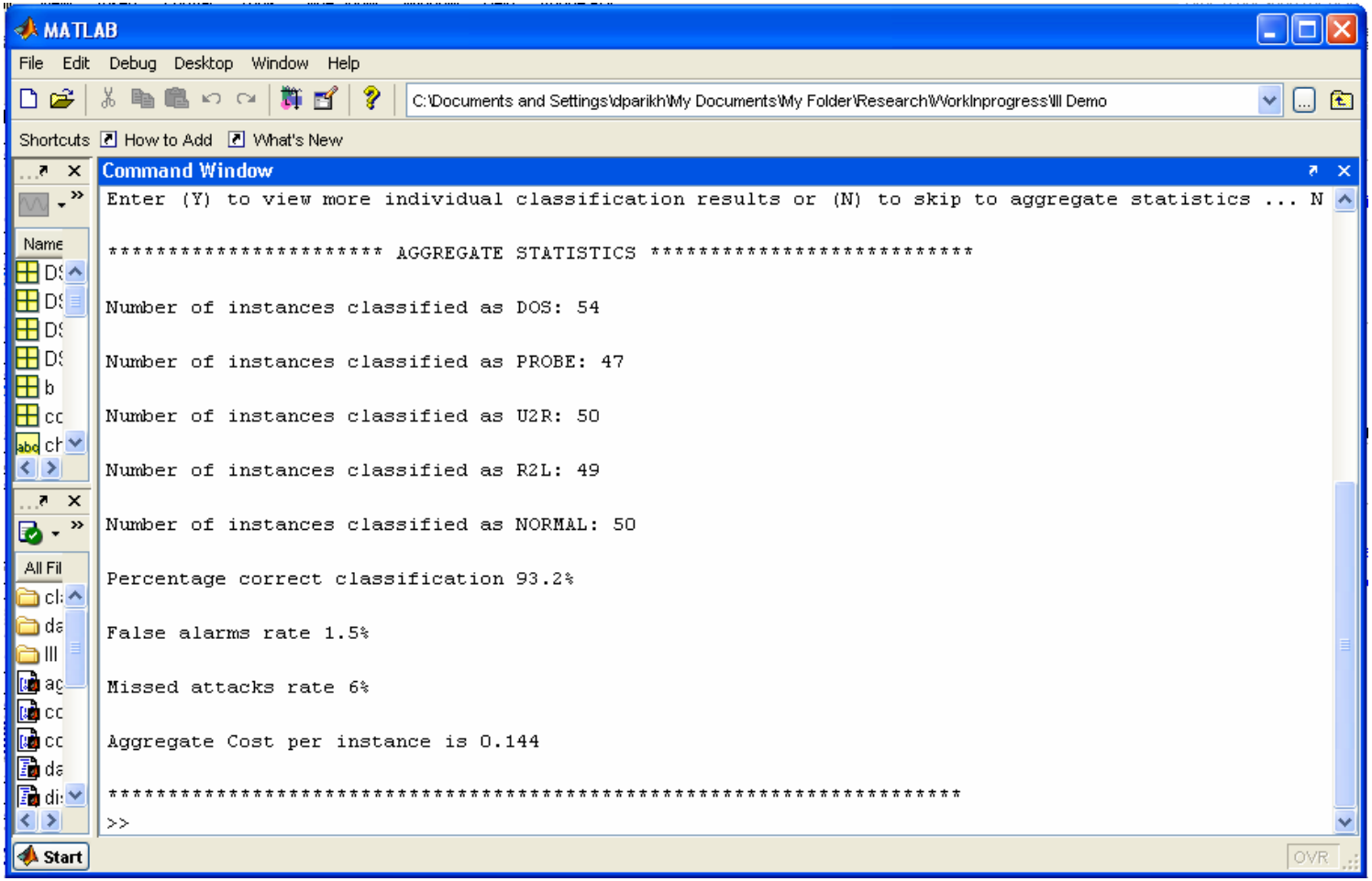
System - Layout

Demo

Conclusion

Questions

Demonstration



Carnegie Mellon

Current Status

- Database analyzed and set up
- Identified most informative sources of information for each attack type
- Working prototype
- Data fusion incorporated
- Comparable classification performance

Future work

- Incorporate adaptability
- Incorporate fine tuning towards minimizing the cost function
- Computationally efficient

Summary

- Pattern recognition techniques for intrusion detection
- Neural networks
- Data fusion
- Adaptability

Questions?

Carnegie Mellon