AI*IA 2003

Fusion of Multiple Pattern Classifiers

PART IV

Applications of Multiple Classifiers Fusion

Multiple classifiers theory and methods can play an important role for all the applications where decision-level fusion is required (MCS Int. Workshop series *www.diee.unica.it/mcs*)

Question raised during Int. School on Ensemble Methods for Learning Machines (Sept. 2002):

Identify an application/problem for which fusion of multiple classifiers surely does not work (i.e., does not provide any benefit)

No application was identified clearly !

F. Roli's position: for cases where multiple classifiers cannot improve performance, they can anyway increase reliability. So, the use of MCS it is just a matter of cost vs. benefit.

Applications of Multiple Classifiers Fusion

Among the others:

- Remote Sensing
- Biometrics
- Documents Analysis and OCR
- Data Mining and KDD

During the MCS workshops, the current and future impact of multiple classifiers on applications was discussed in round tables *(www.diee.unica.it/mcs)*

➢ In the following, I give some selected examples, focusing on security applications (biometric for personal identification, intrusion detection in computer networks)

Fingerprint Classification

Task: to distinguish among five fingerprint classes. Classification is the processing step performed before fingerprint identification to speed up the process.

(a)

(b)

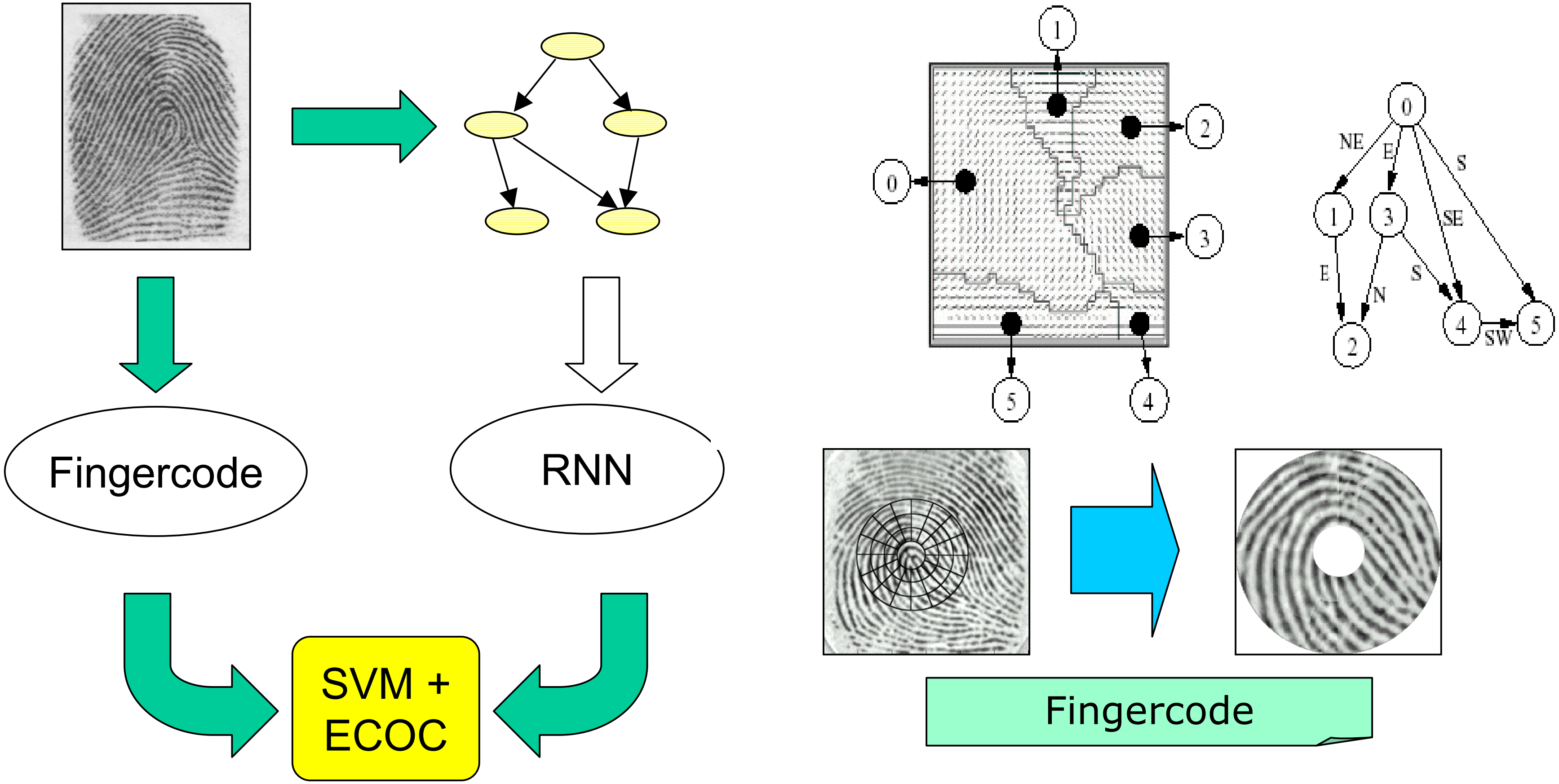
(c)

(d)

(e)

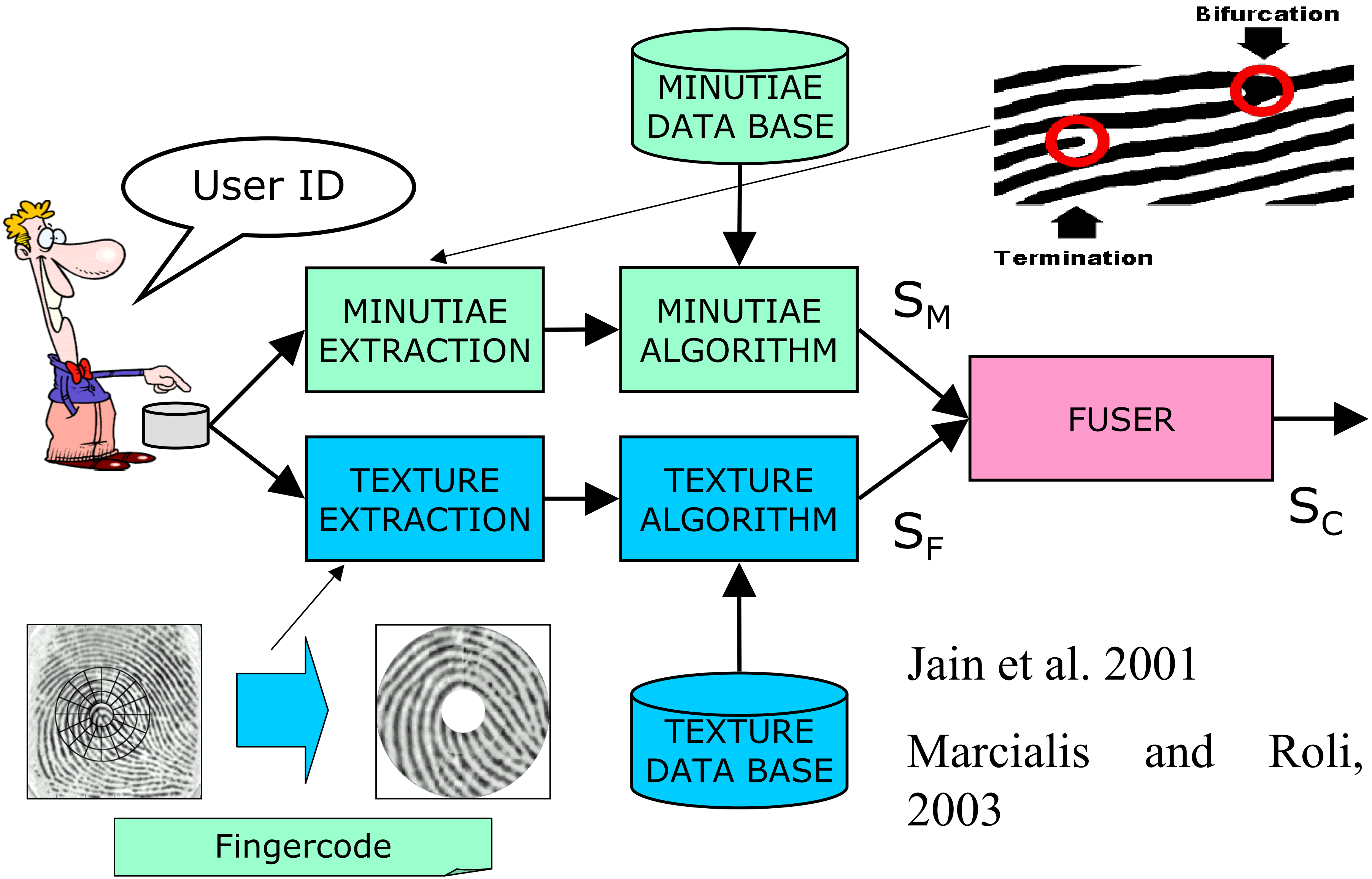
(a) Whorl (W)
(b) Right Loop (R)
(c) Left Loop (L),
(d) Arch (A)
(e) Tented Arch (T)

Fusion of structural and statistical fingerprint classifiers



Marcialis, Roli, and Serrau (2003) reported performance improvement over the best individual classifier

Fingerprint Verification



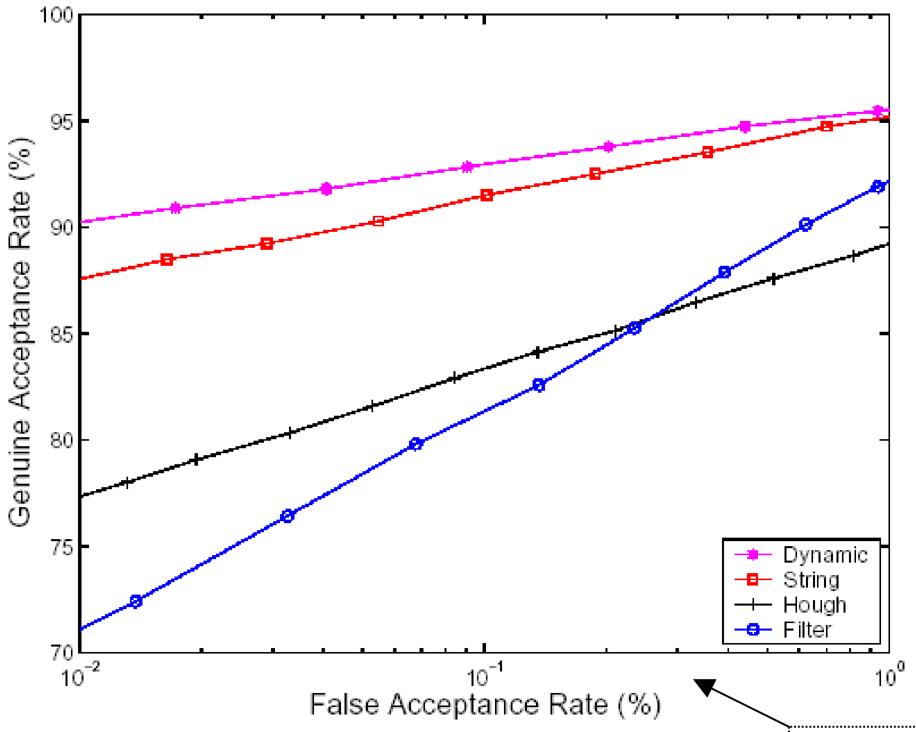
Jain et al. 2001

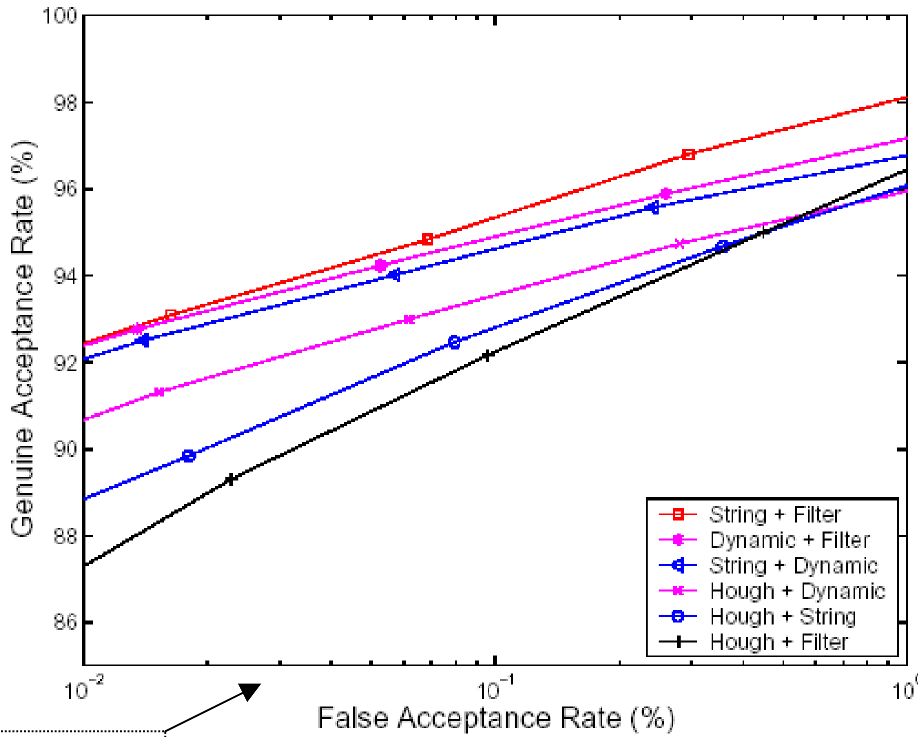
Marcialis and Roli, 2003

Fingerprint Verification

Four different fingerprint matchers were used and combined on a real data base (S.Prabhakar and A.K. Jain, 2001, Marcialis and Roli, 2003).

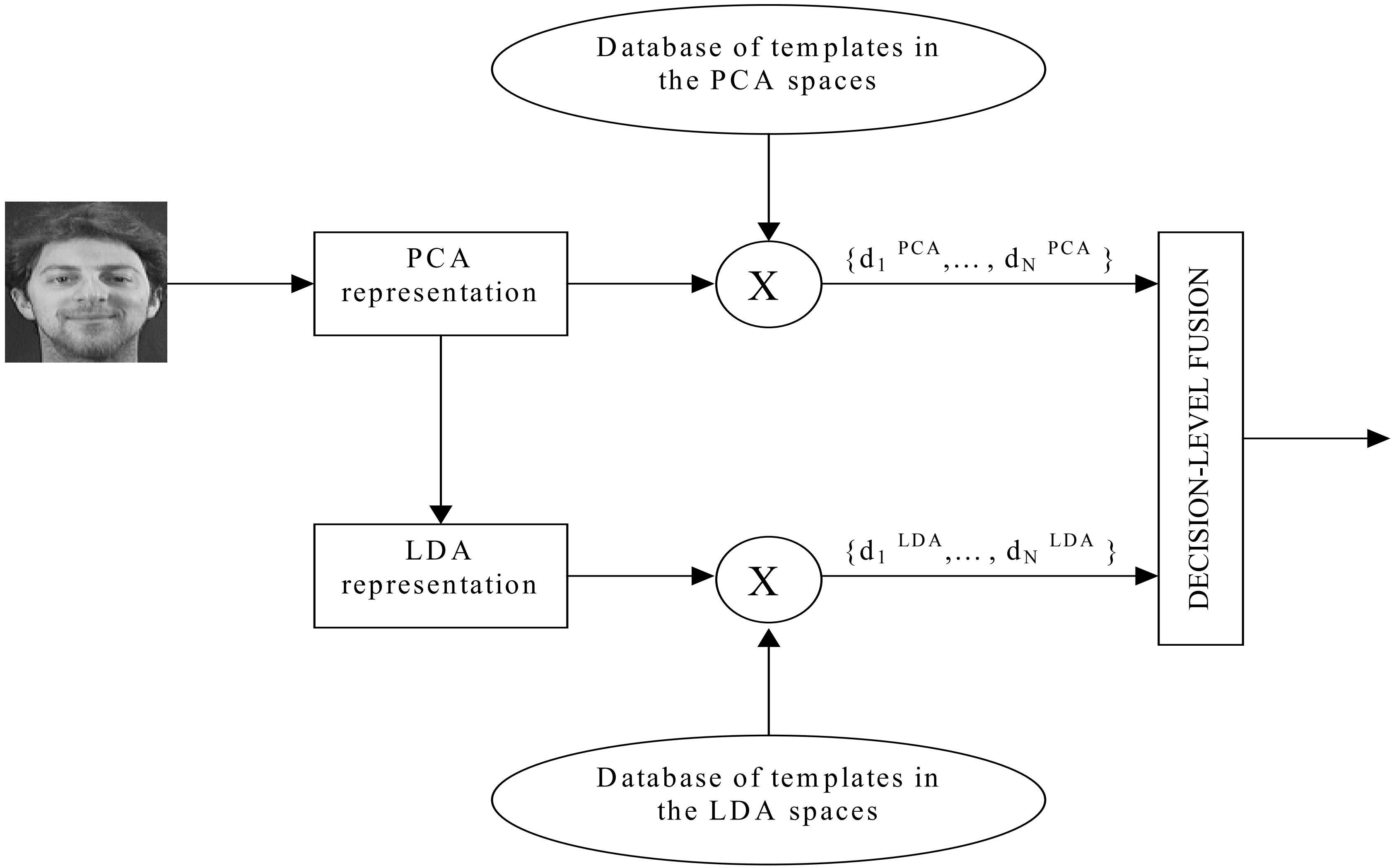
Fingerprint matchers fusion allows improving the performances of the individual matchers (improvement of 3%-5%).



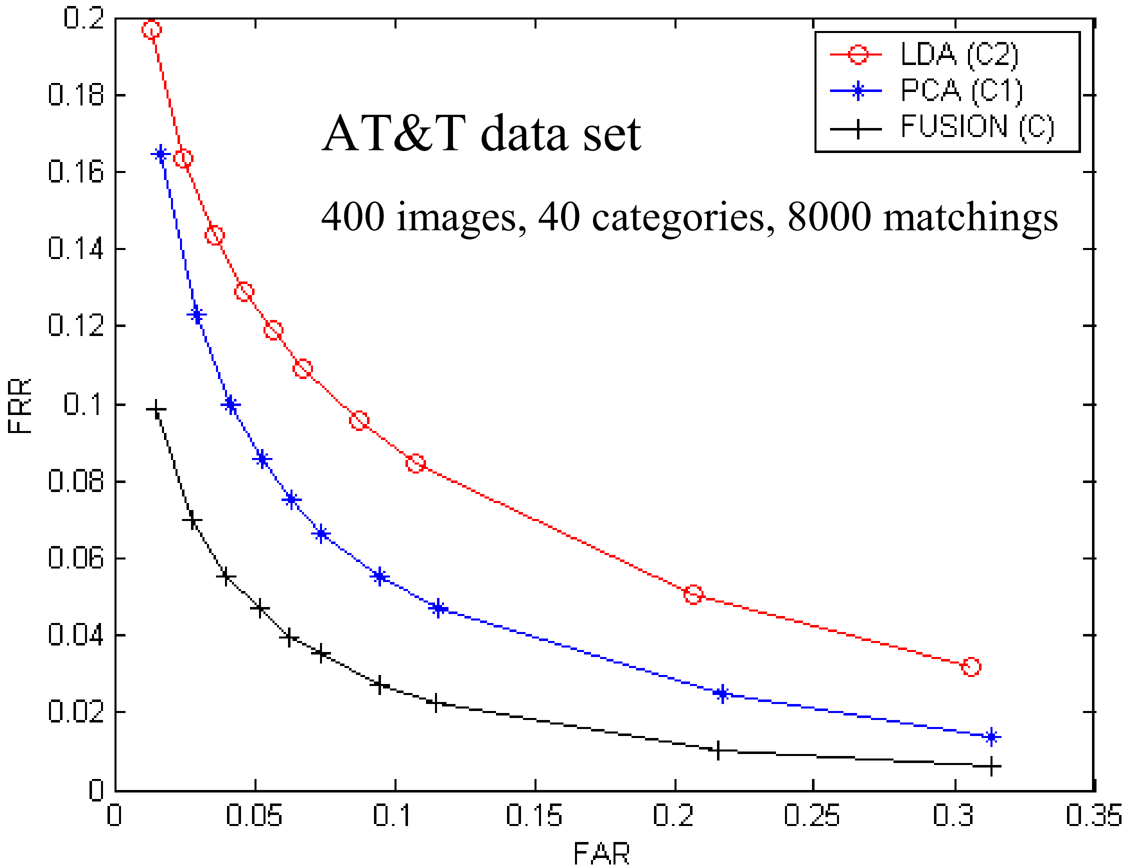


ROC curves

Face Verification



Fusion for Face Verification



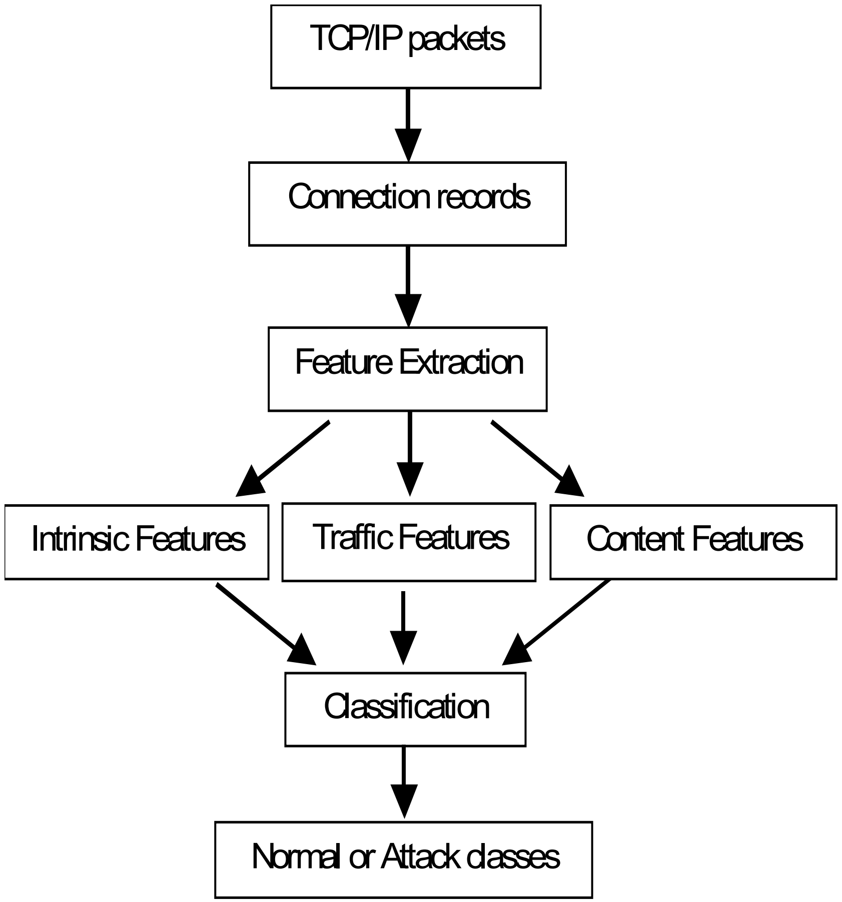
Fusion is based on simple fixed rules (Marcialis and Roli, 2002).

ROC curves are shown (False Rejection Rate as a function of the False Acceptance Rate).

Fusion allows improving the performance of the individual algorithms.

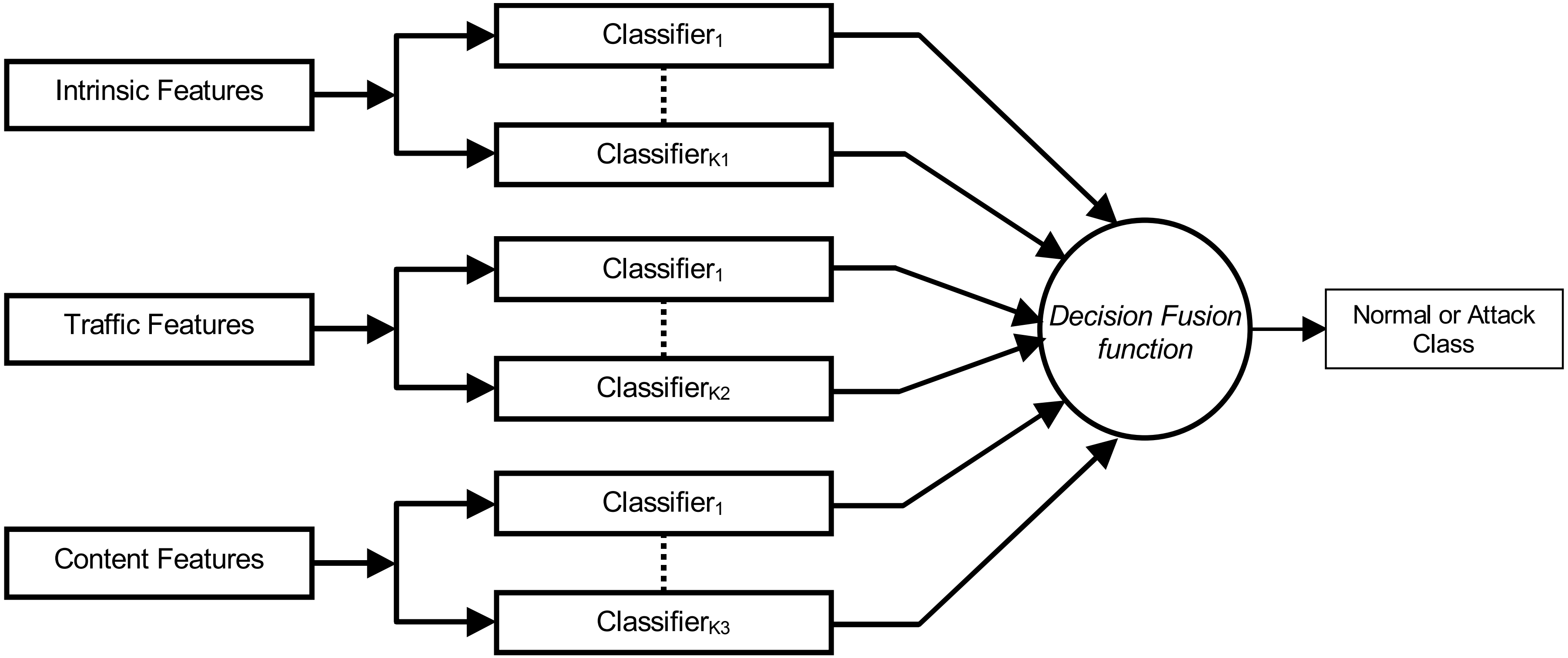
Intrusion Detection in Computer Networks

Roli and Giacinto, 2002



Intrusion Detection formulated as a Pattern Recognition Problem

Intrusion Detection in Computer Networks



- Giacinto and Roli (2002) results showed that MCS can improve detection performances

- In particular, performances related to "unknown" attacks and false alarm rate

MCS for High Reliability Pattern Classification

- The previous examples of applications showed that MCS can improve the reliability (expressed in terms of error-reject trade-off, FAR vs. FRR, etc.) over the individual classifiers

  - For linear combiners, Roli and Fumera (2002, 2003) showed that these experimental evidences have theoretical support.
  - They showed that linear combination of classifiers outputs can improve, under specific assumptions, the error-reject trade-off of individual classifiers

State of the Art

- After more than a decade of research in the MCS field, where do we stand ?

•What are the main results ?

•Today what a designer of pattern classification system has in her/his hands that was not available ten years ago ?

State of the Art

- After more than a decade of research, we have two main approaches to pattern classifiers fusion (T.K. Ho):

Coverage optimisation methods: a simple combination function is given. The goal is to create a set of complementary classifiers that can be combined optimally

-Bagging, Random Subspace, etc.

Decision optimisation methods: a set of carefully designed and optimised classifiers is given and unchangeable, the goal is to optimise the combination function

-Abstract-level/Rank-level/Measurement-level,Fixed vs. Trained, Parallel vs. Others, etc.

State of the Art

- For both coverage and decision optimisation methods, we have many empirical evidences of their effectiveness

•Some theoretical explanations of their effectiveness, theoretical supports

•However, complete theories are a matter of on-going research. An unifying framework is clearly beyond the state of the art

➢For a given task, the choice of the most appropriate combination method lies on the "old" paradigm of model evaluation and selection !!

State of the Art: The Gloomy Side

- But if the choice of the most appropriate combination method lies again on the "old" paradigm of model evaluation and selection

•Then, one could say that MCS research simply moved the original problem to a different level

•Instead of looking for the best classifier, now we look for the best combination rule

Are we fallen in a vicious circle ? Soon are we looking for the best set of combination rules ? (T.K. Ho; F.Roli, 2002)

State of the Art: Good News

Fusion of "naturally" distinct classifiers

In some applications, the fusion of classifiers arise naturally from the application context:

- Multi-modal biometrics where we have to fuse distinct classifiers that work on distinct inputs (face, speech, fingerprints)
- Fusion of algorithms based on distinct computational mechanisms
- etc.

- For such cases, MCS research (in particular, decision optimisation methods) provided a large "bag" of fusers

State of the Art: Good News

Difficult Pattern Classification Tasks

For difficult tasks, MCS research provided alternative approaches:

•For tasks where the optimisation of a single algorithm is difficult, time consuming, and tedious, coverage optimisation methods can be a solution. Generating and fusing multiple "weak" classifiers can be simpler !

•For tasks where a set of "diverse" classifiers are available (e.g., based on different computational methods), decision optimisation methods can provide further performance improvements.

State of the Art: Good News

•MCS provides solutions for cases where the selection of the best classifiers is very difficult, as the estimates of generalization error are optimistically biased ("apparent" error)

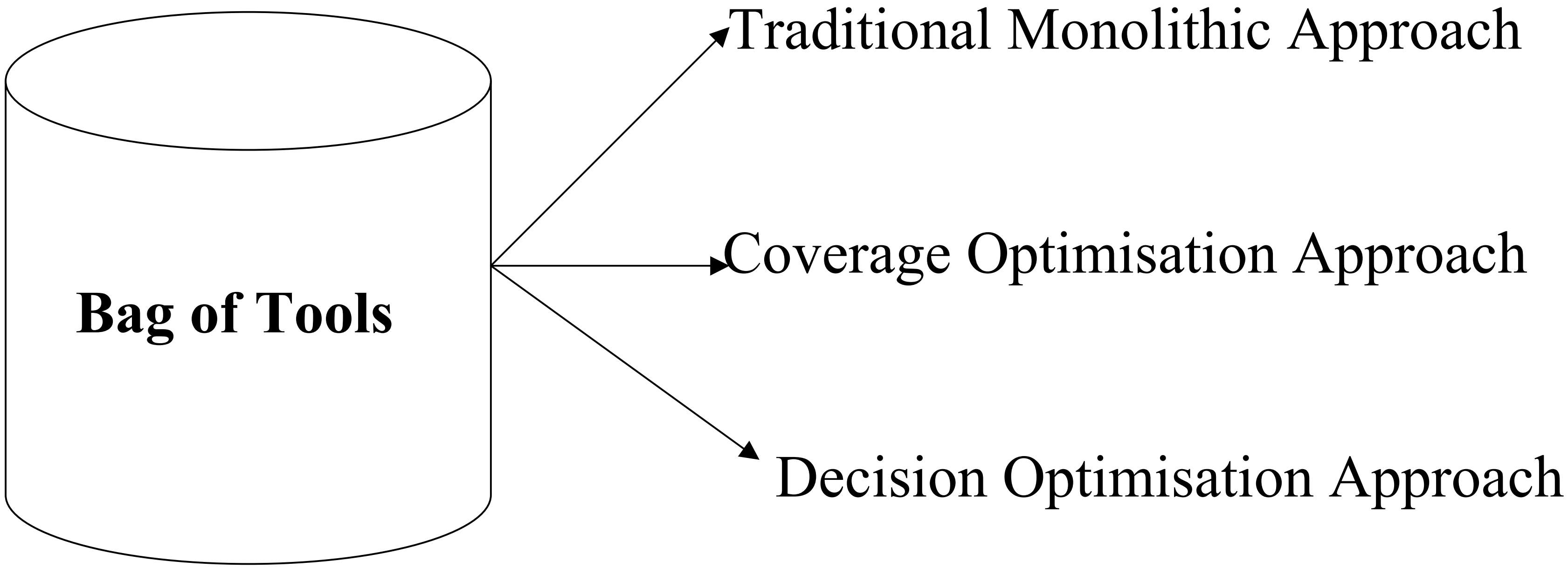
•MCS provides solutions for cases where the optimisation of an individual classifier is very difficult. Increase of design effort provide very small improvements. In such cases, coverage optimisation methods can be a solution

•MCS can avoid the choice of arbitrary initial conditions, or the tuning of difficult design parameters

•MCS provides solutions for cases where a single classifier that is optimal on the whole feature space does not exists! Different classifiers can be optimal in different regions of the feature space ("Modular" MCS).

State of the Art: Good News

The New Scenario



MCS research provided new “tools” to exploit better the Kanal’s bag of tools (1974)

Future Work

•Theoretical side:

-Complete theories for coverage and decision optimisation approaches, and unifying frameworks

-Rigorous explanations of the operation of main combination algorithms

•Practical side

-Practical guidelines for choosing the most appropriate single or combined system for the task at hand

Room for Optimism

As while theoretical, experimental, and engineering research make progress, I think that:

•The number of practical tasks for which we will able to choose (in a probabilistic sense) the best single or combined system will increase

•For each task, the size of the set of single or combined systems that we will have to evaluate and compare for selecting the best one will decrease