Institute for Infocomm Research
A*STAR

JOENSUUN YLIOPISTO

NIST 2006 Speaker Recognition Evaluation

Institute for Infocomm Research (IIR) & University of Joensuu (JoY) Joint Submission

Presented by:

Prof. Pasi Fränti

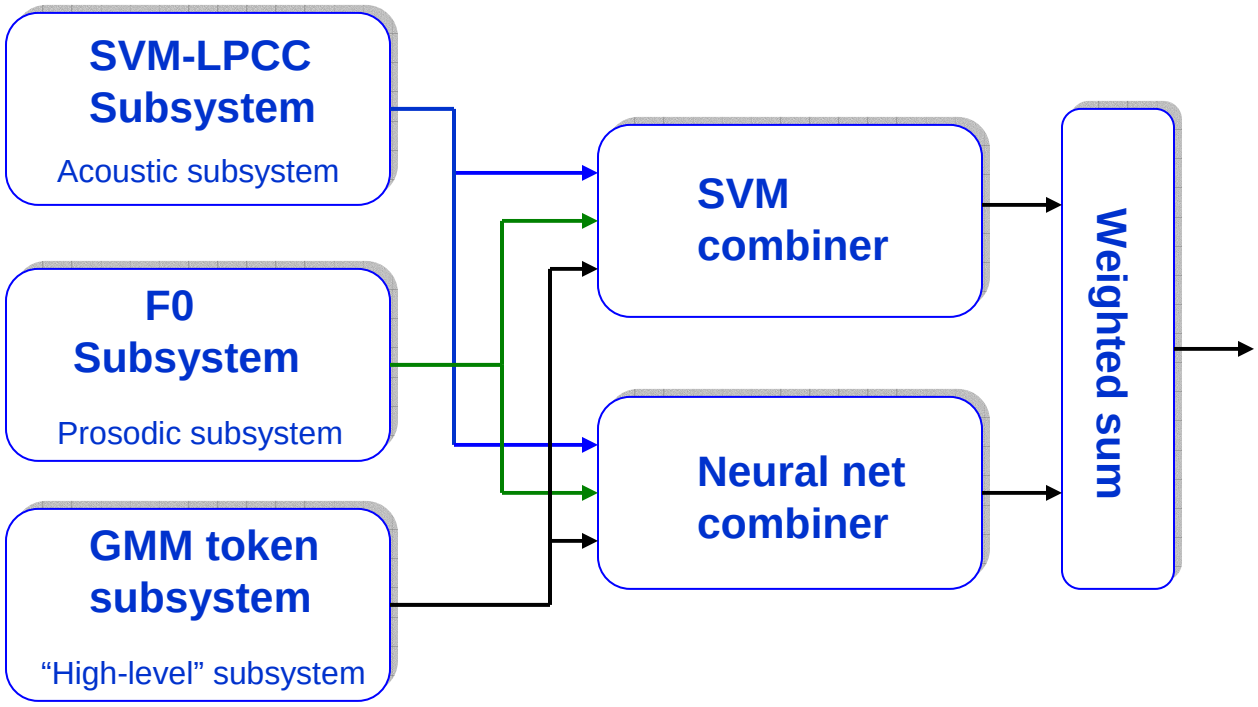
Speech and Image Processing Unit
Department of Computer Science
University of Joensuu, Finland

IIR-J Submissions

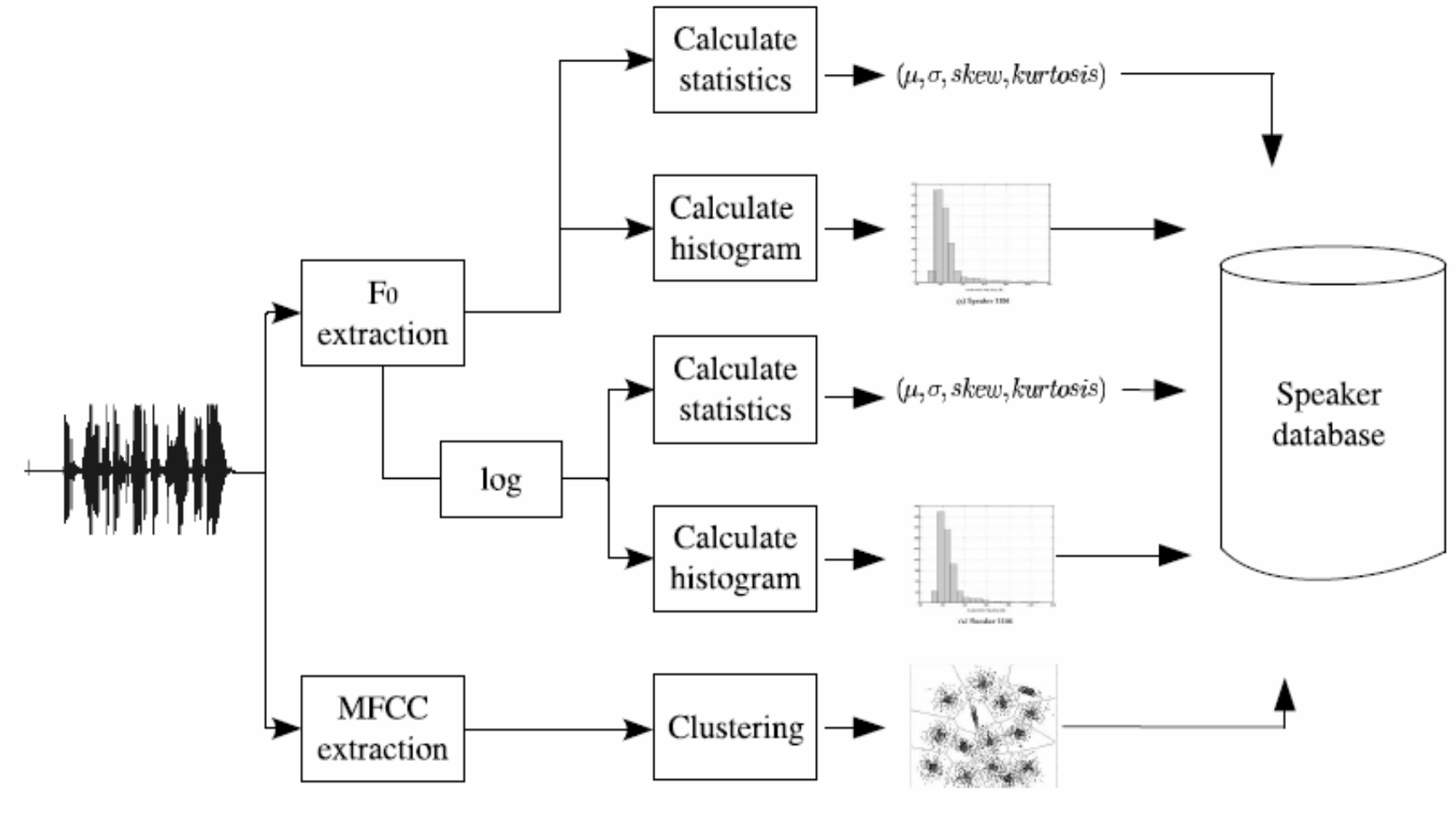
Same as the IIR submissions

		Test segment condition			
		10sec 2chan	**1conv 2chan**	**1conv Summed chan**	**1 conv aux mic**
Training condition	**10sec 2chan**	10sec4w-10sec4w			
	1conv 2chan	1conv4w-10sec4w	1conv4w-1conv4w		
	3conv 2chan	3conv4w-10sec4w	3conv4w-1conv4w		
	8conv 2chan	8conv4w-10sec4w	8conv4w-1conv4w		
	3conv summed chan				

System Diagram

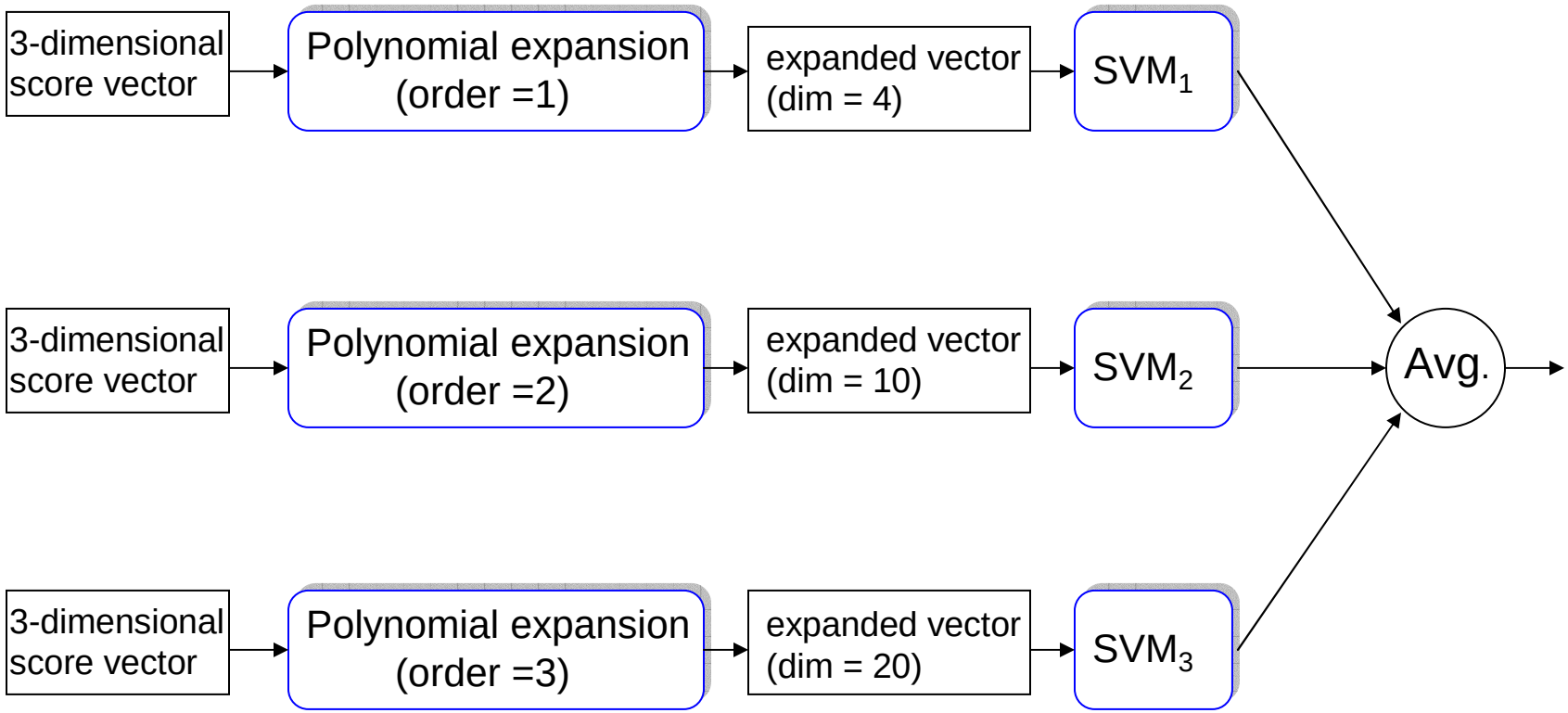


F0 Subsystem



SVM Combiner

Same as in IIR submissions



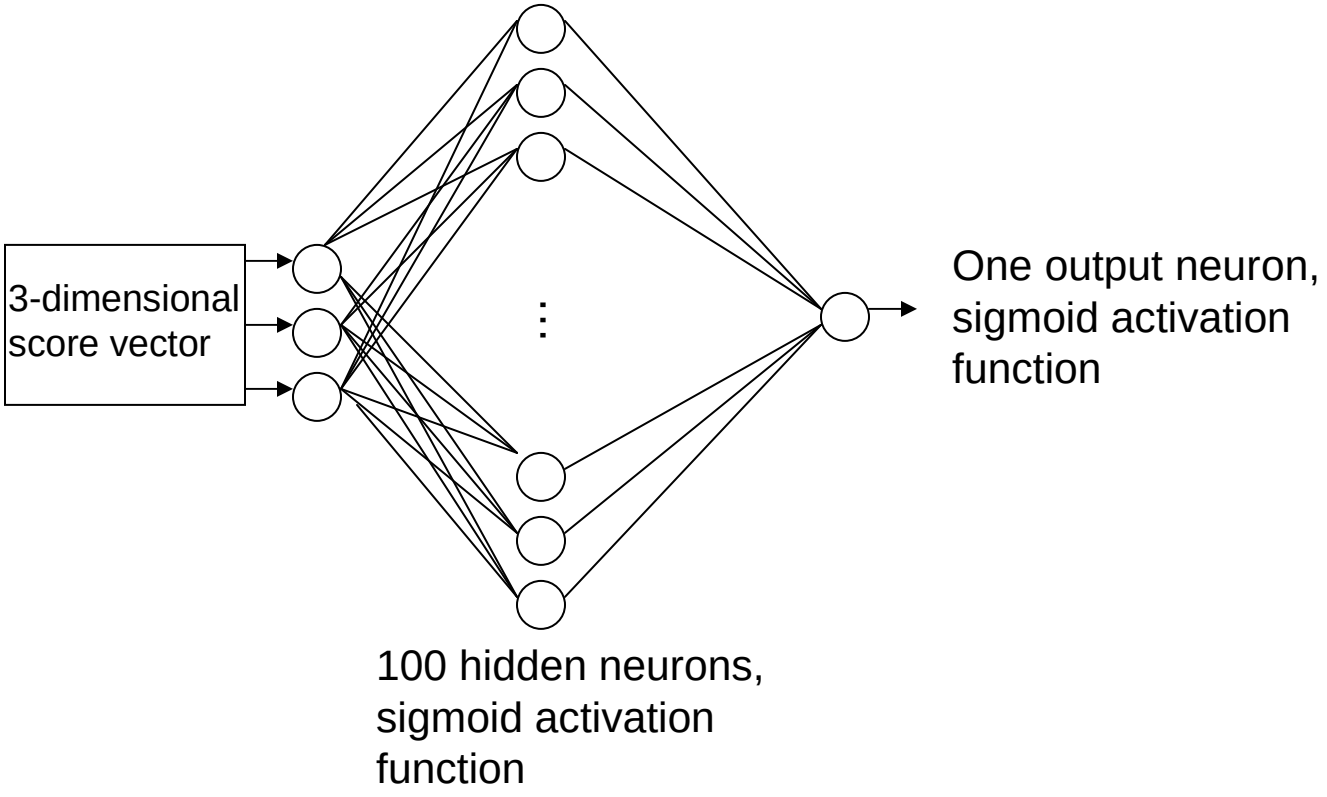
Polynomial expansion (order = 3)

$$(x_1, x_2, x_3) \mapsto (1, x_1, x_2, x_3,$$
$$x_1^2, x_2^2, x_3^2, x_1x_2, x_1x_3, x_2x_3,$$
$$x_1^3, x_2^3, x_3^3, x_1x_2^2, x_1x_3^2, x_1^2x_2, x_1^2x_3, x_2^2x_3, x_2x_3^2, x_1x_2x_3)$$

SVM training:

- Score vectors labeled “+1” for genuine speakers and “-1” for impostors
- Linear SVM, trained using SVMTorch package

Neural net combiner



Training:

- Desired output: 1 for genuine speakers and 0 for impostors
- Scaled conjugate gradient algorithm, with Netlab toolbox for Matlab

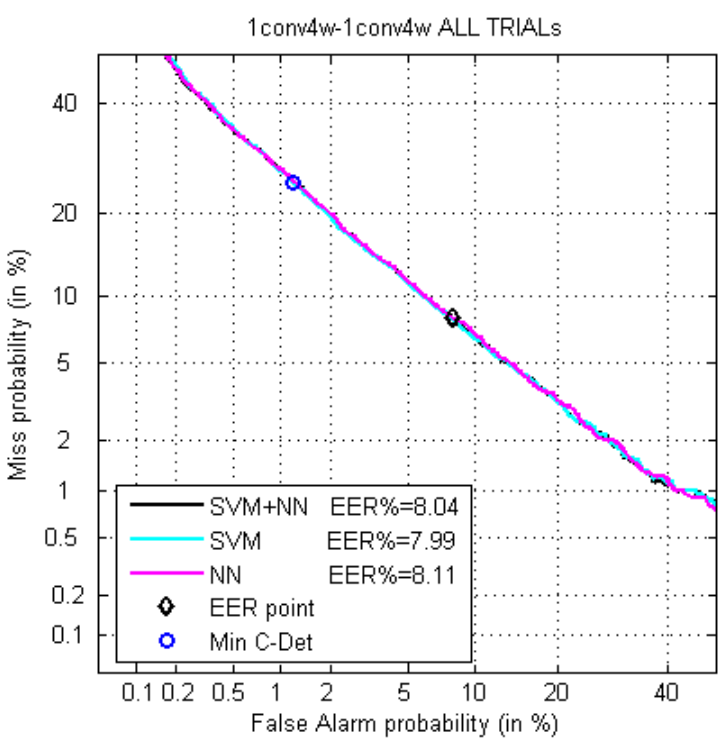
Training of Classifiers

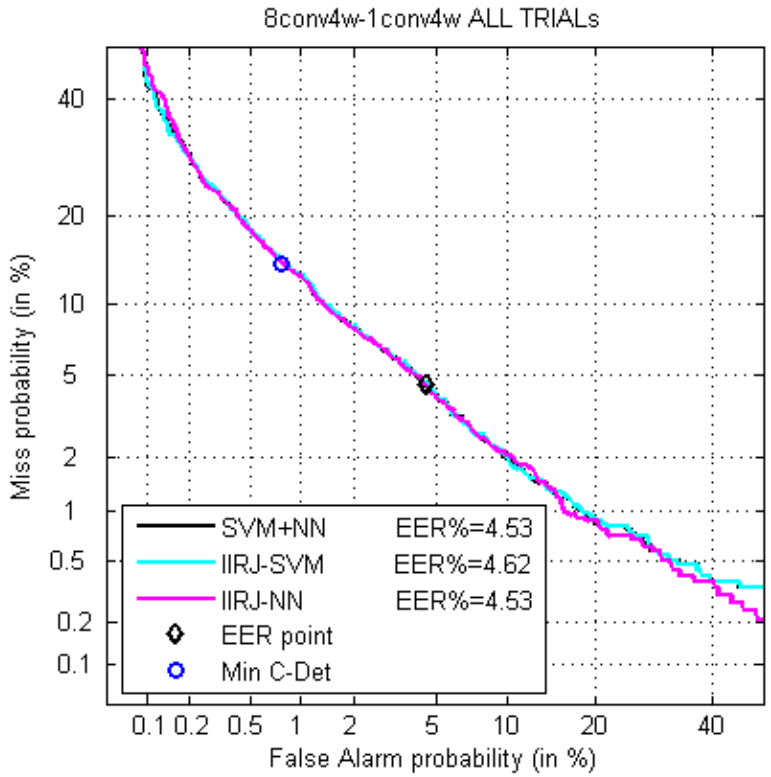
	SVM-LPCC	F0	GMM token
Tokenizer training data	N.A.		NIST SRE 2002 corpus
Background (UBM) training data	2 gender-dependent background data sw3p1, sw3p2, sw2p2 and sw2p3	N.A.	NIST SRE 2004 corpus
Cohort data (Tnorm)	Evaluation set of the NIST SRE 2004 corpus	N.A.	
Dev.data (train/test)	NIST SRE 2005 corpus		
1conv4w-1conv4w EER %	11.02	39.26	18.67

IIR-J Submissions

- Three submissions, difference in the score combiner:
  - Submission 1 (**primary**): NN+SVM combiner
  - Submission 2: NN combiner
  - Submission 3: SVM combiner
- Differences with the IIR submission:
  - Only three subsystems instead of six
  - Inclusion of F0 subsystem
  - Different fusion strategy

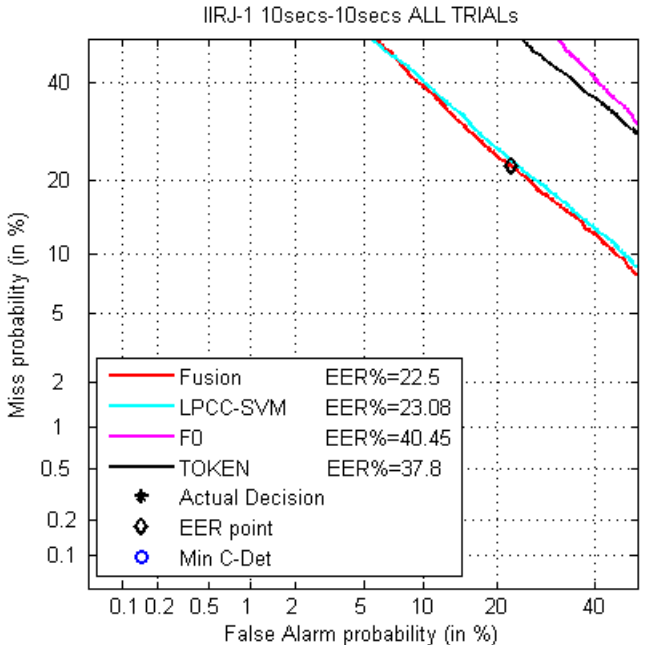
Effect of Combiners



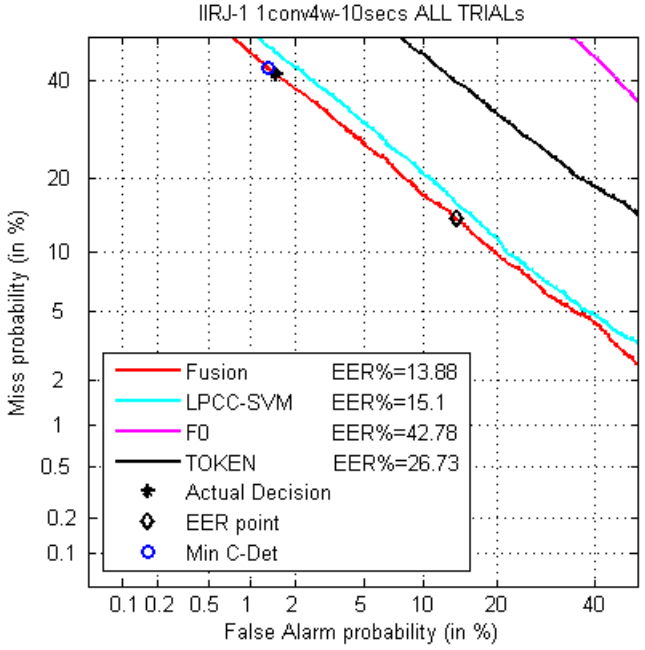


Subsystems (primary submission)

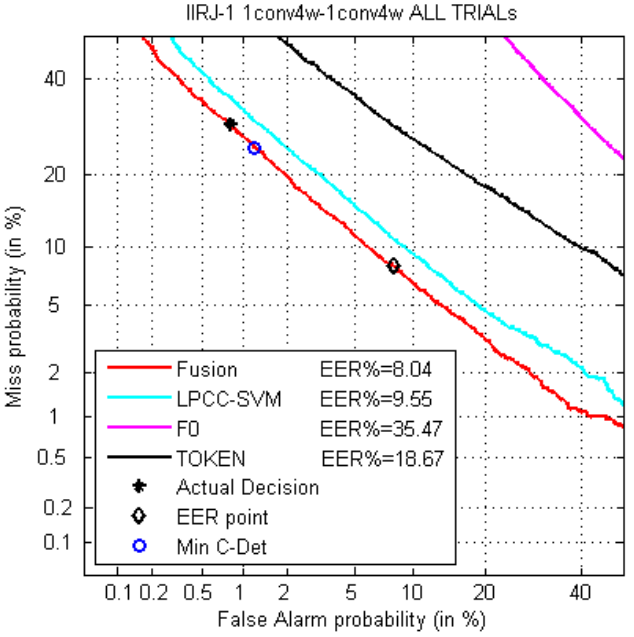
10sec-10sec



1conv-10sec

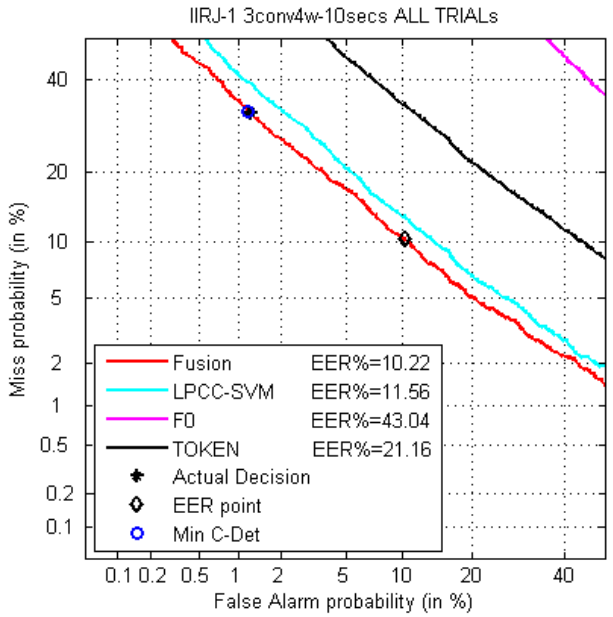


1conv-1conv

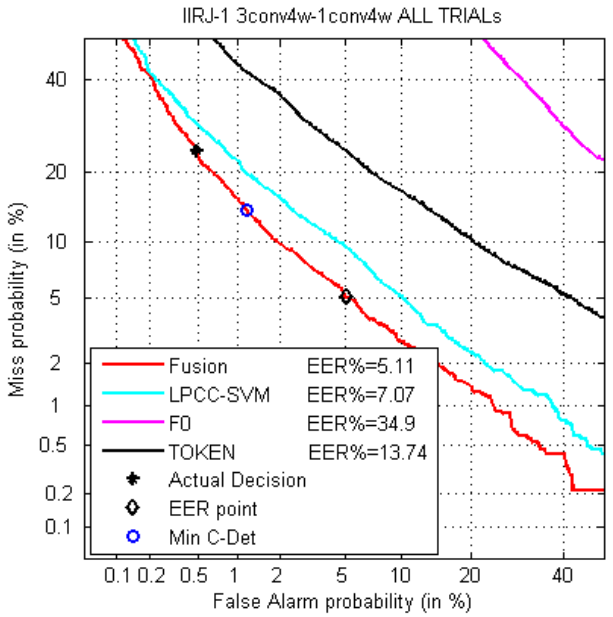


Subsystems (cont.)

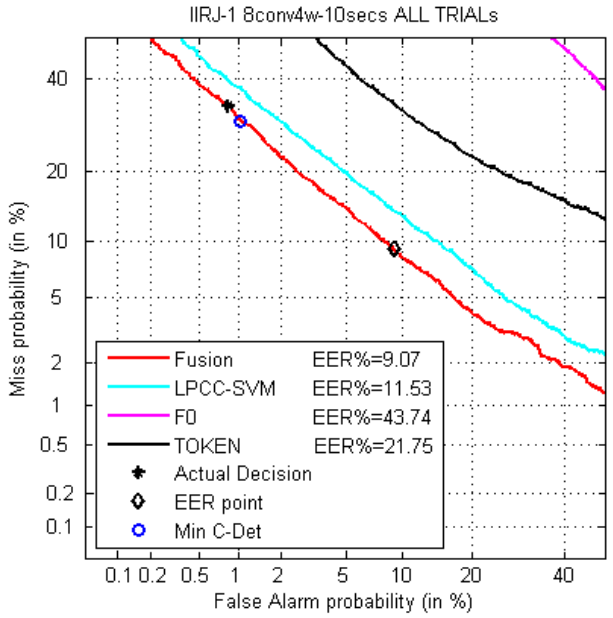
3conv-10sec



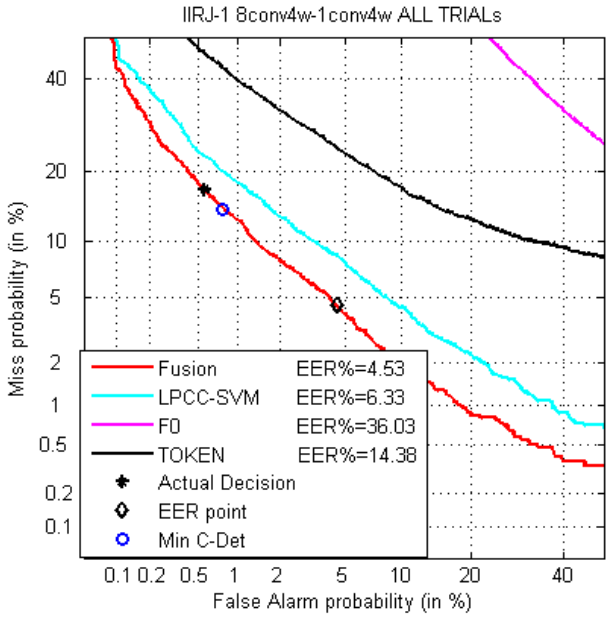
3conv-1conv



8conv-10sec



8conv-1conv



Comparison of IIR and IIRJ Primary Systems