Aristotle University of Thessaloniki
Polytechnic School
Department of Electrical and Computer Engineering
Intelligent Systems and Software Engineering Labgroup

Continuous implicit authentication of smartphone users by navigational and behavioral data

Christos Emmanouil

Professor Supervisor: **Andreas Symeonidis**

PhD Supervisor: **Thomas Karanikiotis**

Contents

Motivation

Constantly increasing number of smartphone users

Storage of personal and business data

Need to seccure the data stored on these devices.

Concerns about the adequacy of existing authentication methods.

Need to implement new authentication methodologies.

Continuous – Implicit Authentication

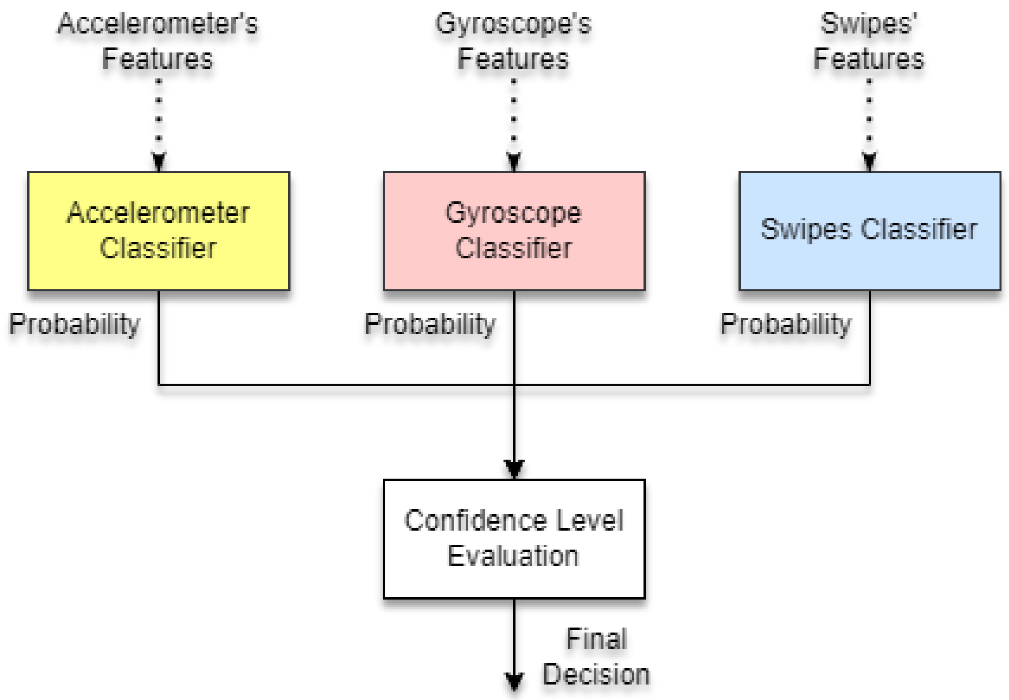
Advantages :

- Enhanced security system
- Better user experience
- Ability to use behavioral characteristics
  - Easy adjustment
  - Low implementation cost
  - Development prospects

Problems :

- High sampling rates
- High power consumption
- Secondary devices (wearables)
- Insufficient evaluation
  - Small amount of data
  - 'lab' data
  - 'Wrong ' metrics
- Insufficient data during execution

Main idea



Objective :

- Satisfactory levels of security and transparency
- Use of data generated by the smartphone
- Tolerant of errors and/or missing data

Questions :

- Data set
- Feature extraction and preprocessing
- Structure of classifiers
- Trust subsystem structure
- Objective evaluation

Dataset

BrainRun :

- Set of behavioral data
- Motion and gesture sensor data
- Data collection application (android & iOS)
- 5 different games, with different levels of difficulty

Characteristics :

- 2218 users
- 60% male, 26% female, 14% unknown
- 90% Android , 10% iOS

Games & Final Sets
(after applying selection criteria) :

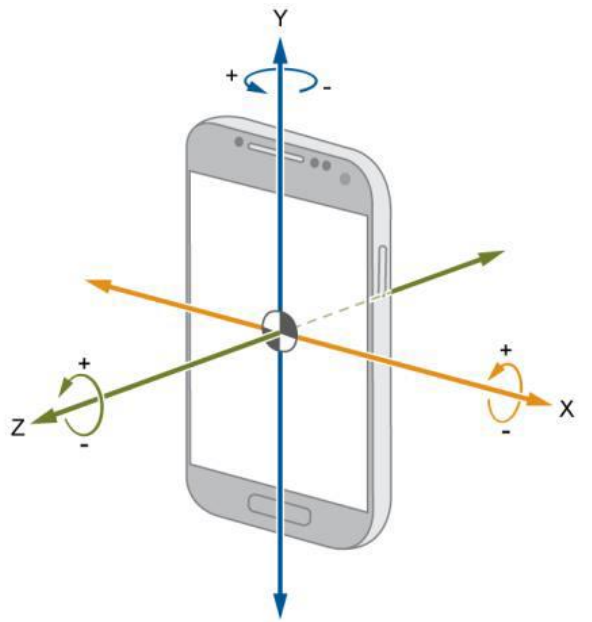
Games	Data type	Number of Training Users	Number of Evaluation Users
Mathis	Acc, Gyr, Swp	15	24
Focus	Acc, Gyr, Swp	15	30
Reacton	Acc, Gyr, Swp, Tap	15	45
Memory	Acc, Gyr, Tap	15	44
Speedy	Acc, Gyr, Tap	15	45

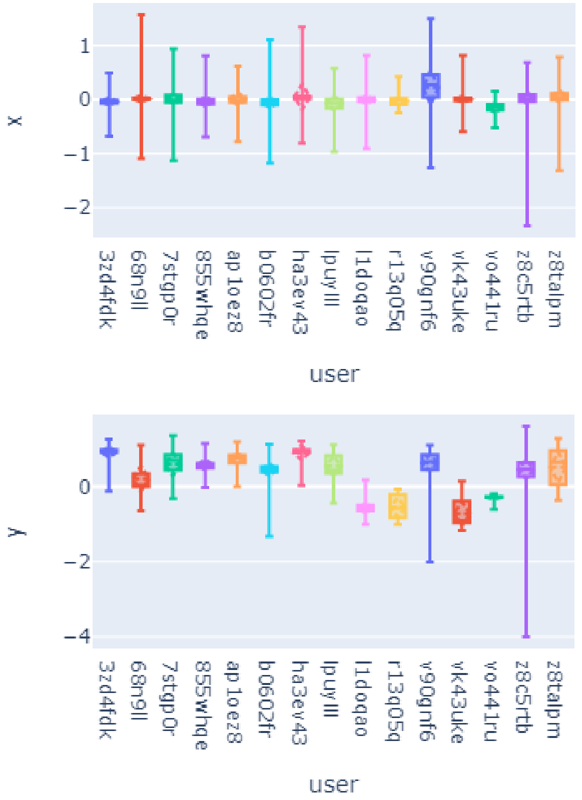
Acc: Accelerometer, Gyr: Gyroscope, Swp: Swipe

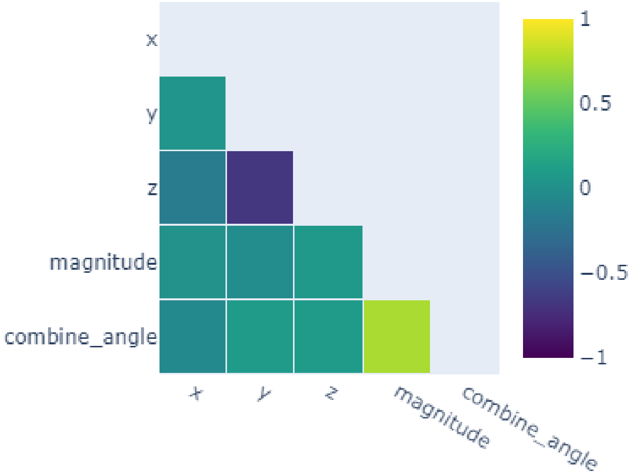
Methodology

Extract Features

Accelerometer, Gyroscope (1)





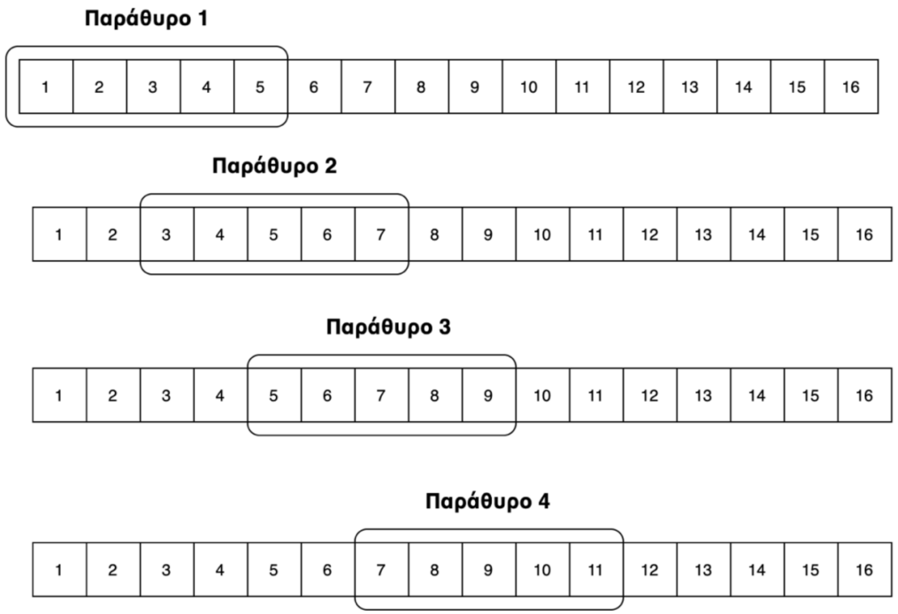


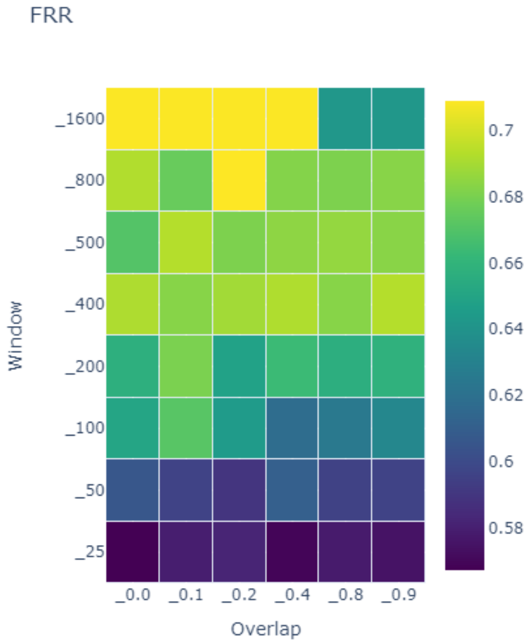
Selected:
x, y and magnitude

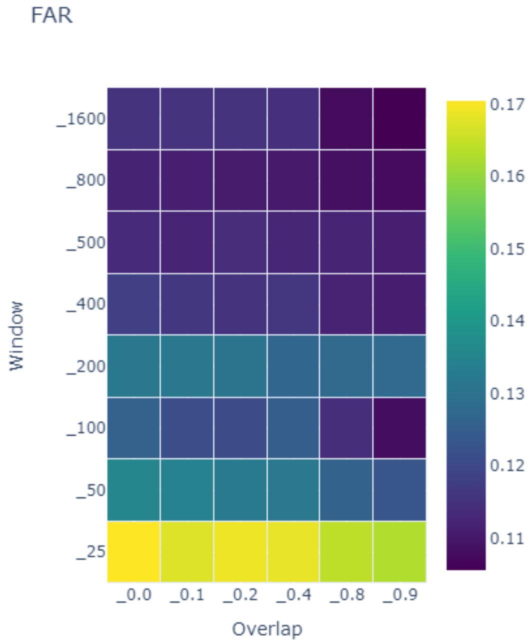
Methodology

Extract Features

Accelerometer, Gyroscope (2)





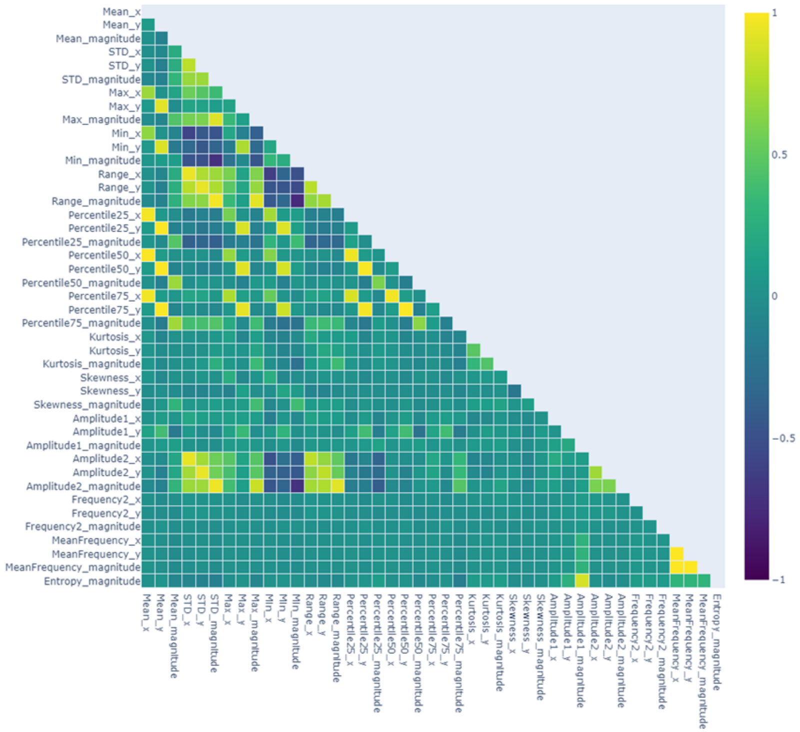


Selected:
Window Size: 50 samples
Overlap: 60%

Methodology

Extract Features

Accelerometer, Gyroscope (3)

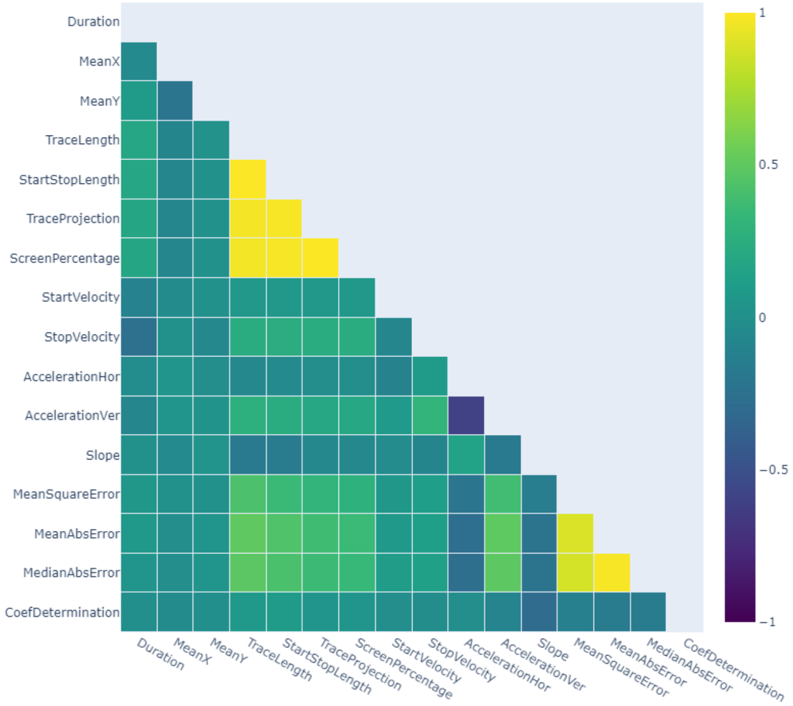


Sensos	Features L0	Features L1
Accelerometer	x	Mean, STD, Max, Min, Percentile25, Percentile50, Percentile75, Kurtosis, Skewness, Amplitude1, Amplitude2, Frequency2, Mean Frequency
	y	Mean, STD, Max, Min, Percentile25, Percentile50, Percentile75, Kurtosis, Skewness, Amplitude1, Frequency2
	magnitude	Mean, STD, Max, Min, Percentile25, Percentile50, Percentile75, Kurtosis, Skewness, Amplitude, Frequency2
Gyroscope	x	Mean, Max, Min, Percentile75, Kurtosis, Skewness, Amplitude1, Frequency2, Mean Frequency
	y	Mean, Min, Kurtosis, Skewness, Frequency2
	magnitude	Mean, Min, Kurtosis, Skewness, Frequency2

Methodology

Extract Features

Gestures



Gesture	Features Final
Tap	Duration
Swipe	Duration, Mean X, Mean Y, Trace Length, Trace Projection, Start Velocity, Stop Velocity, Horizontal Acceleration, Vertical Acceleration, Slope, Mean Square Error, Coefficient of Determination

Classifiers

What do we know?

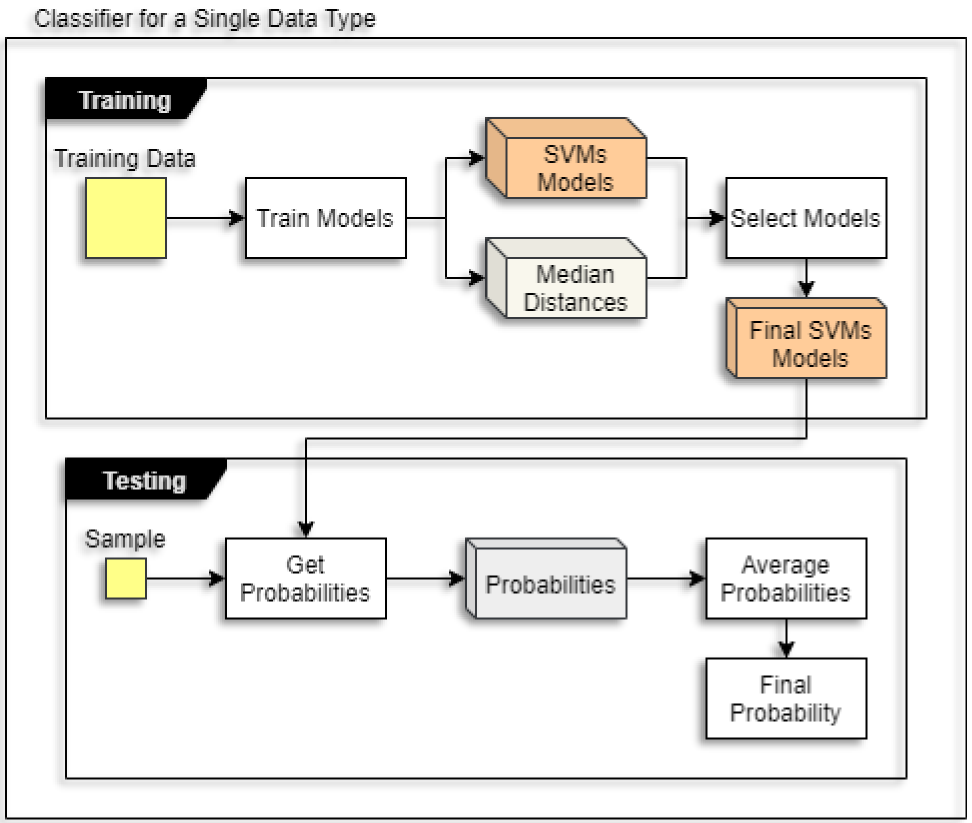
- Single class classification problem
- Solving with RBF-OCSVM
- Impossible to use one model per classifier
- The parameters (nu, gamma) affect the RBF-OCSVMs

What do we recommend?

- Using multiple RBF-OCSVMs , per classifier
- Use a range of values for the parameters
- Collective final decision

Questions :

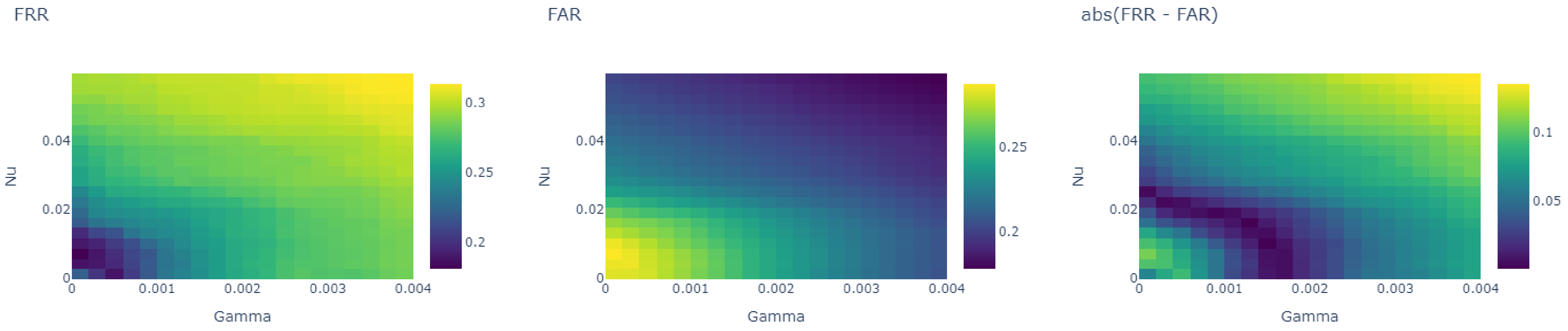
- Range of parameters
- Number of deciding models



Methodology

Classifiers

Parameters Range

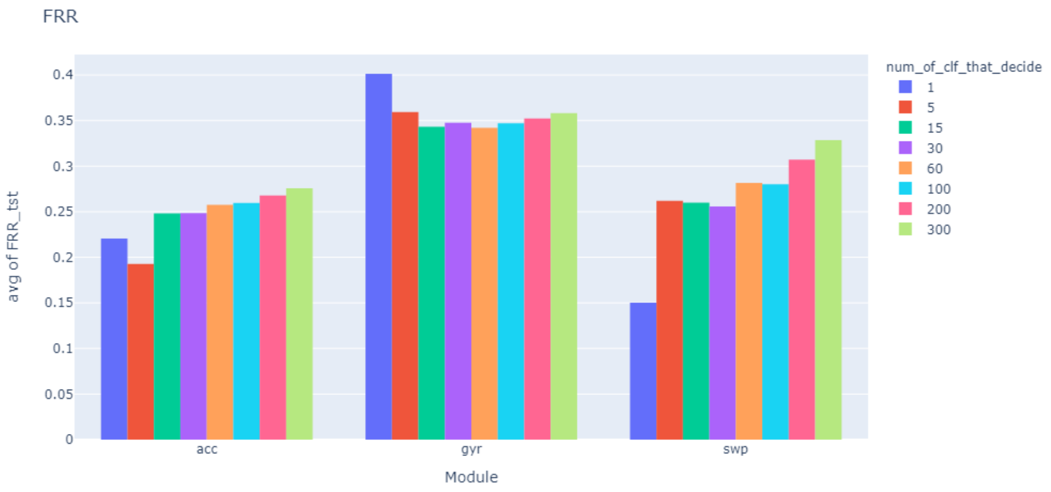


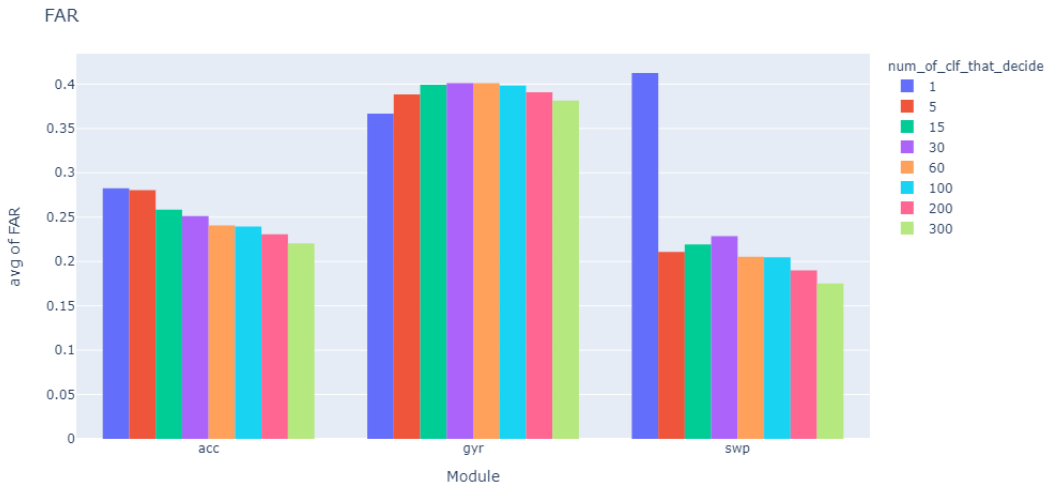
Type	Nu			Gamma		
	Start Value	End Value	Step	Start Value	End Value	Step
Accelerometer	0.001	0.06	0.003	0.0001	0.004	0.0002
Gyroscope	0.11	0.31	0.01	0.001	0.04	0.002
Swipes	0.01	0.21	0.01	0.001	0.06	0.003
Taps	0.02	0.6	0.03	0.7	0.795	0.005

Methodology

Classifiers

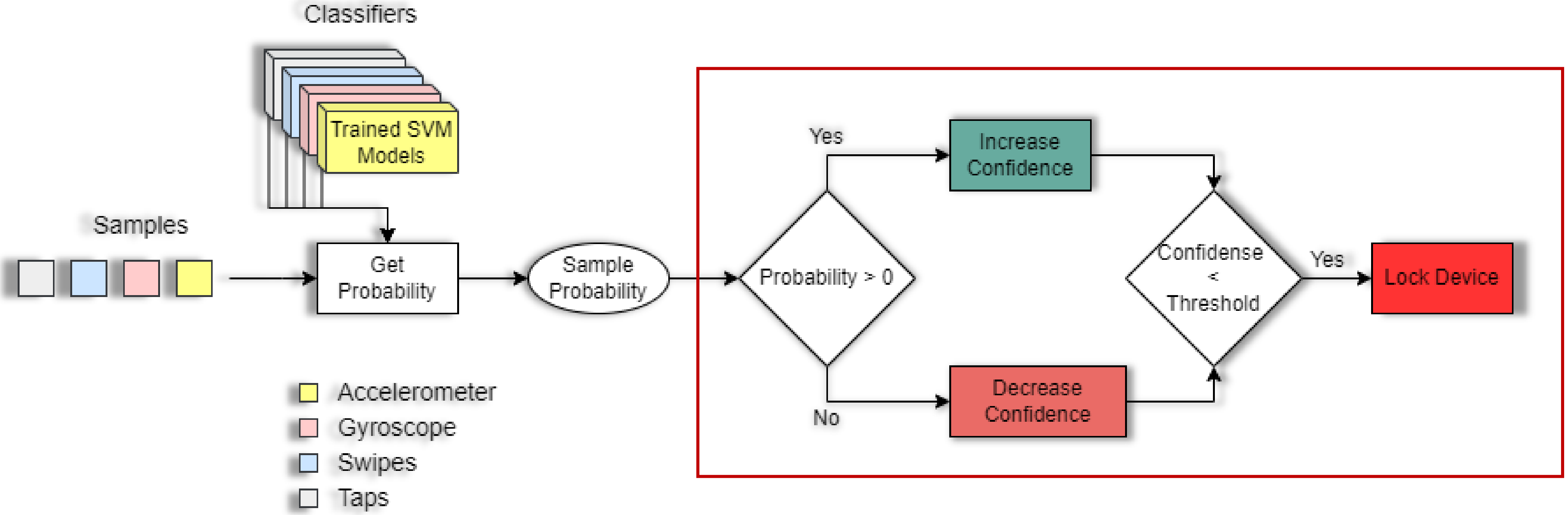
Number of Models





Data Type	Optimal Number
Accelerometer	30
Gyroscope	60
Swipes	60
Taps	60

Confidence Subsystem



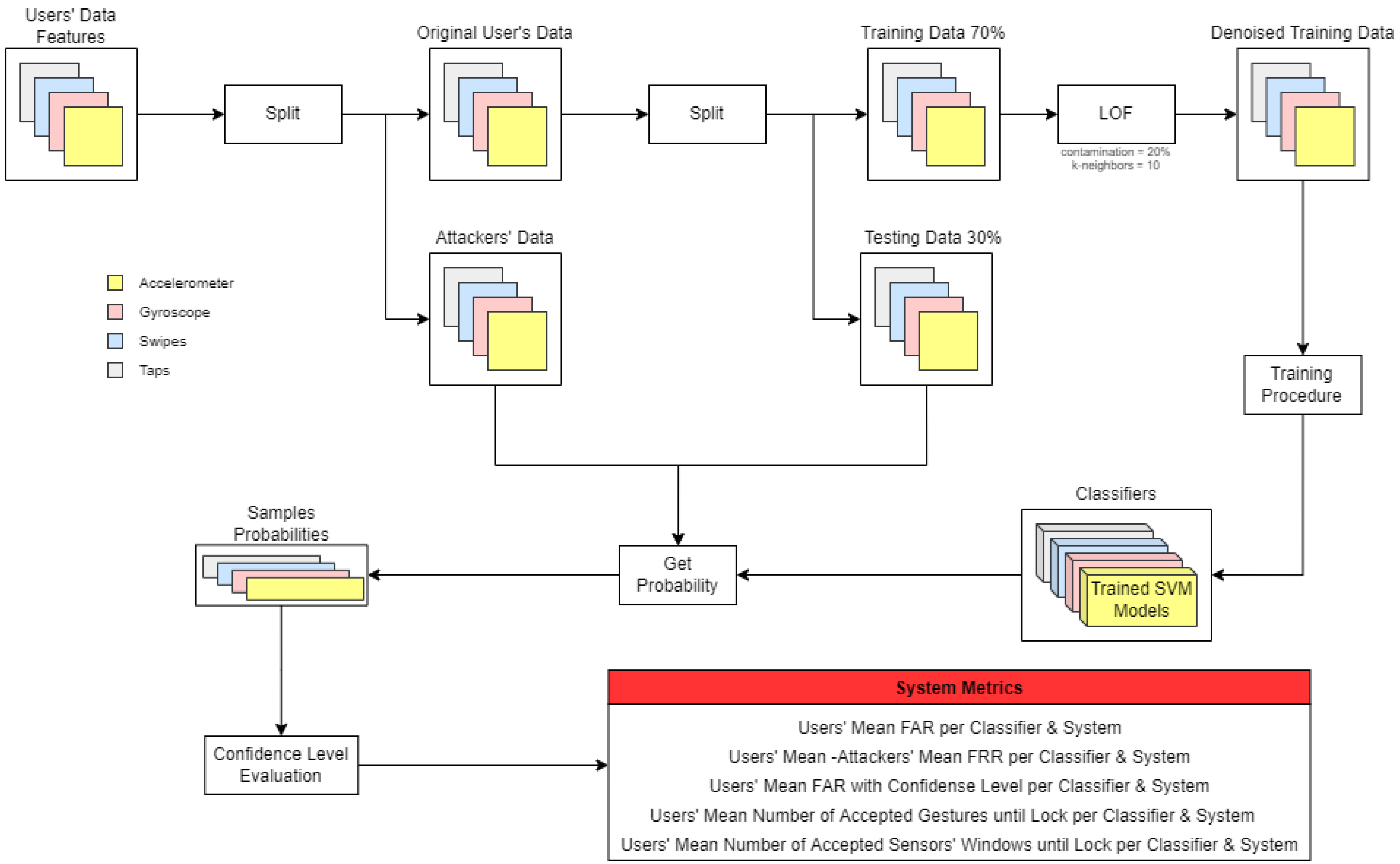
$$CL_n = \begin{cases} CL_{n-1} + PositiveStep(Game) * Weights(DataType) * abs(p), & p > 0 \\ CL_{n-1} + NegativeStep(Game) * Weights(DataType) * abs(p), & p \le 0 \end{cases}$$

Initial Confidence Level	60
Threshold	35

	Mathisis	Focus	Reacton	Speedy	Memoria
Negative Step	-15	-15	-15	-15	-15
Positive Step	+10	+10	+10	+10	+10

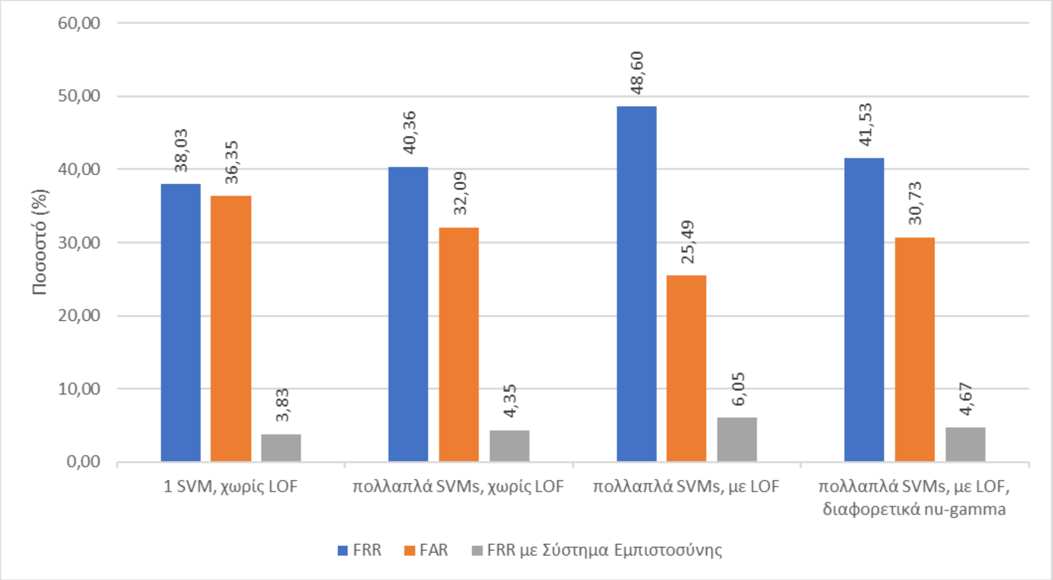
Methodology

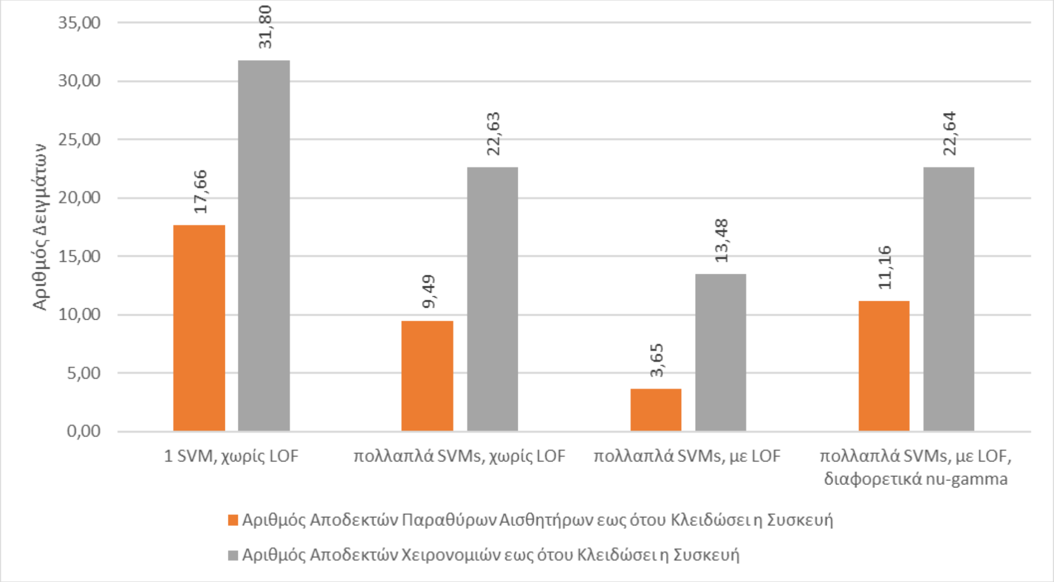
System Summary – Structure of Final Experiments



Results

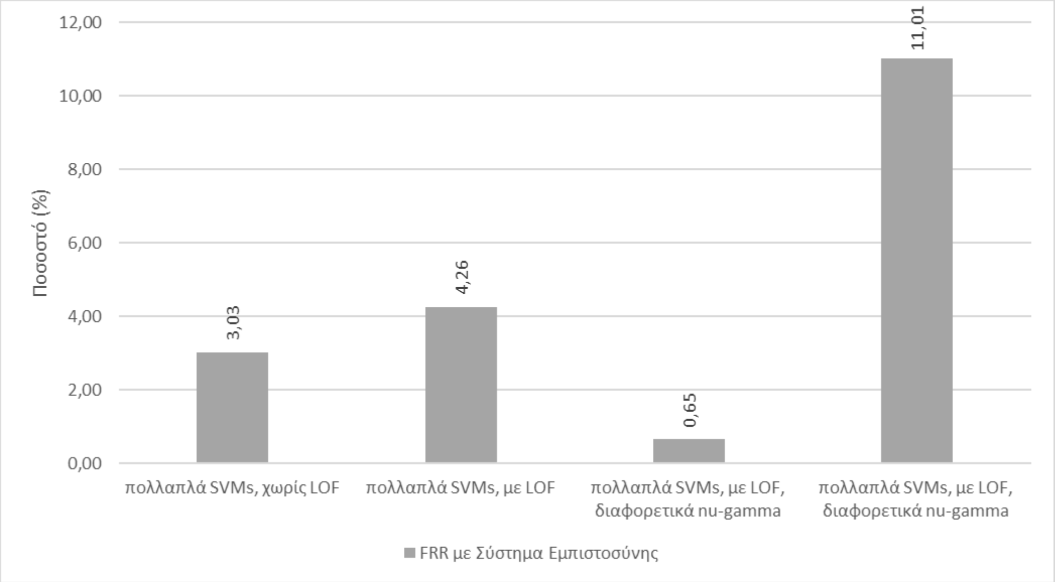
Trust System – Multiple RBF-OCSVMs

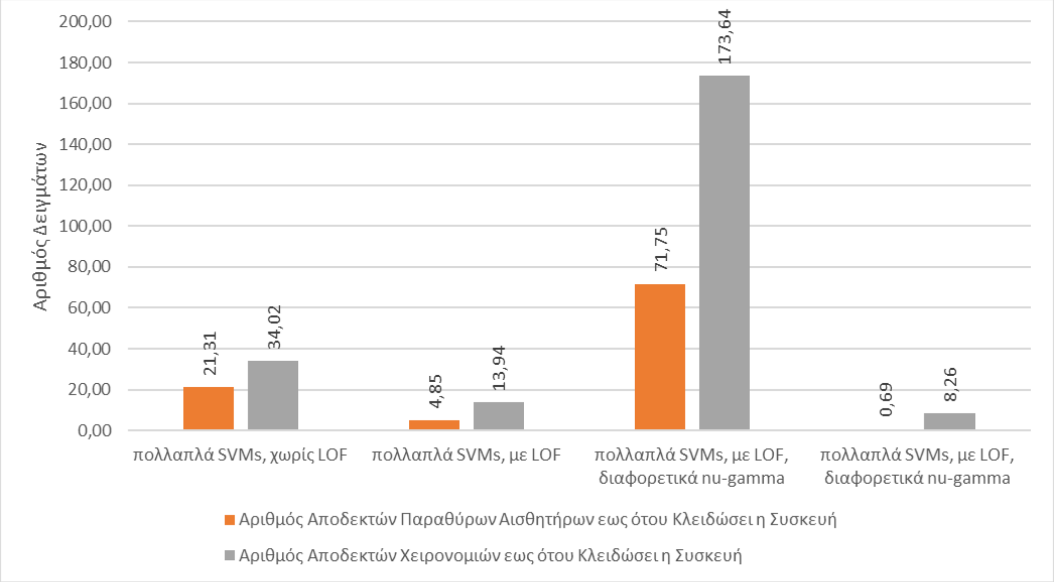




Results

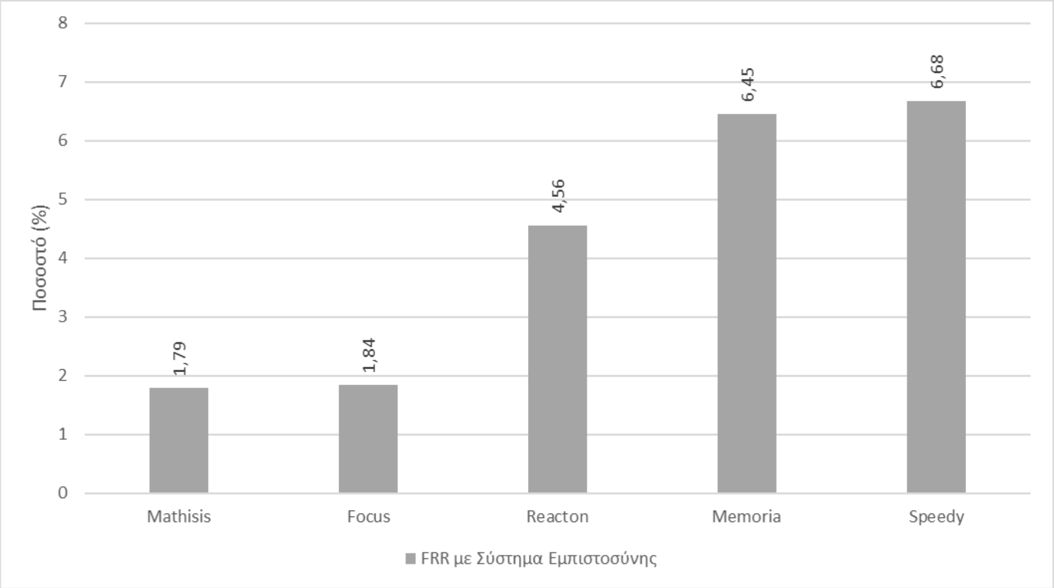
LOF – Nu-Gamma

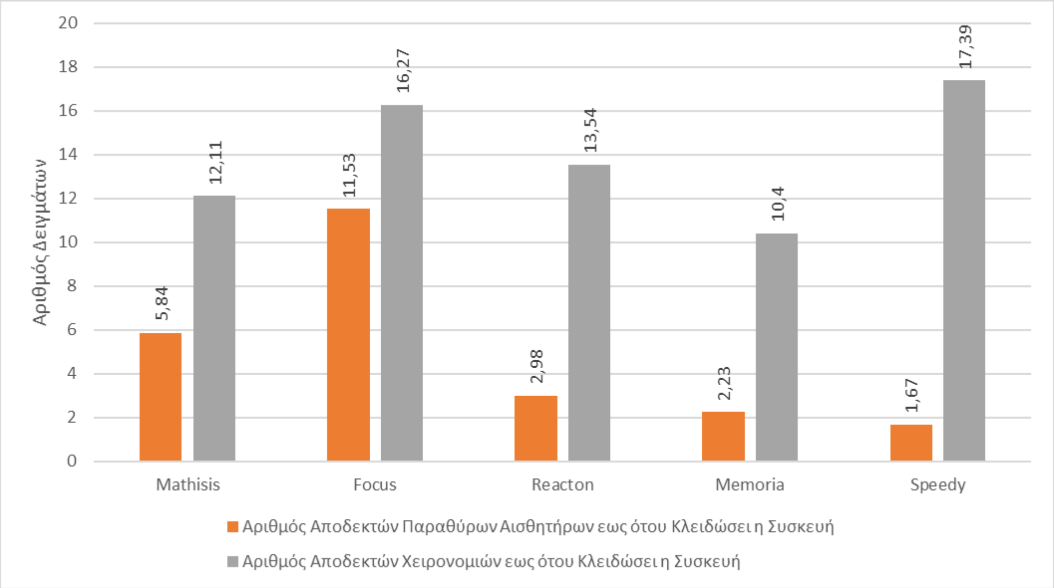




Results

Per Game

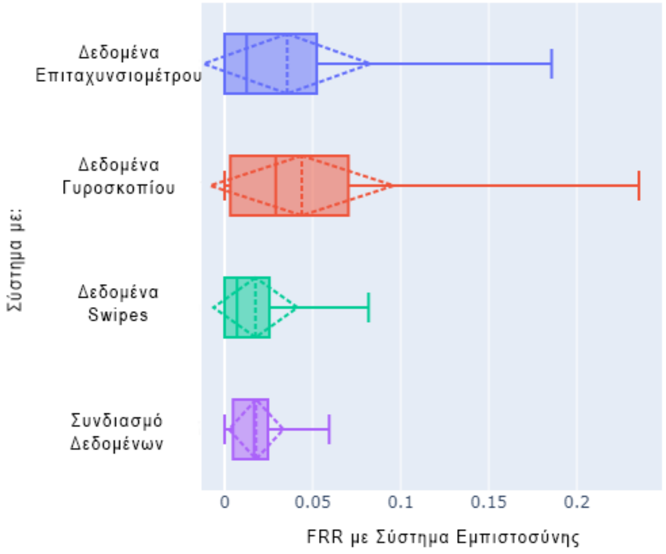


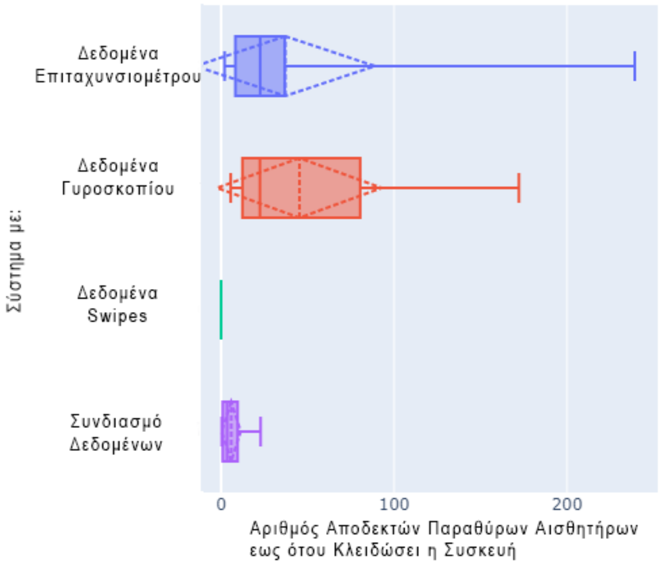


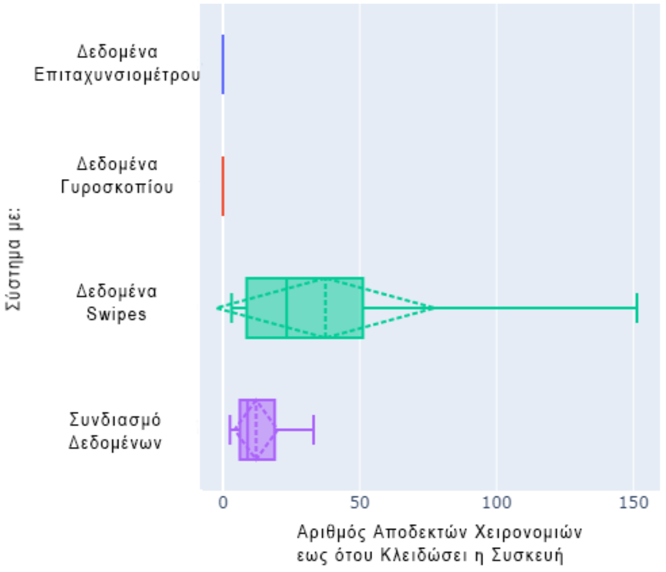
ΑΡΙΣΤΟΤΕΛΕΙΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΘΕΣΣΑΛΟΝΙΚΗΣ

Results

Comparisons (1)







ΑΡΙΣΤΟΤΕΛΕΙΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΘΕΣΣΑΛΟΝΙΚΗΣ

issel Intelligent Systems & Software Engineering Labgroup

Results

Comparisons (2)

	Mathisis	Focus	Reacton	Memoria	Speedy
Σύστημα Αισθητήρων (FRR %)	5,20	6,00	4,30	5,70	5,70
Σύστημα Χειρονομειών (FRR με Σύστημα Εμπιστοσύνης %)	1,92	1,06	2,32 | 3,58 (Swipes | Taps)	3,44	0,065
Τρέχουσα Εργασία (FRR με Σύστημα Εμπιστοσύνης %)	1,79	1,84	4,56	6,45	6,68

	Mathisis	Focus	Reacton	Memoria	Speedy
Σύστημα Αισθητήρων (FAR %)	4,08	3,50	6,90	1,10	5,40
Σύστημα Χειρονομειών (Αριθμός Αποδεκτών Χειρονομειών)	1,70	3,92	8,08 | 11,37 (Swipes | Taps)	21,83	277,47
Προκείμενο Σύστημα (Αριθμός Αποδεκτών Δειγμάτων Αισθητήρων & Χειρονομειών)	5,84 & 12,11	11,53 & 16,27	2,98 & 13,54	2,23 & 10,40	1,67 & 17,39

Sensor Pack Size: 500 counts

Sensor Packet Size: ~50 counts

Conclusions

Methodology & Techniques

- Using multiple RBF- OCSVMs serves system security.
- The trust system helps form an easy-to-use system.
- Denoising the training data with LOF improves security.
- The nu and gamma parameters of RBF- OCSVMs play a decisive role in ensuring a balance between security and usability.

System

- Robust to measurement errors.
- Satisfactory security and transparency metrics.
- Quick check
- Objective evaluation

Future Work

Ideas

- Dynamic weights in classifiers
- Option to select nu-gamma ranges
- Combination with context-aware techniques
- Ability to adapt to changes in owner behavior

ΑΡΙΣΤΟΤΕΛΕΙΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΘΕΣΣΑΛΟΝΙΚΗΣ

issel Intelligent Systems & Software Engineering Labgroup

Thanks!

Anomaly Detection – One Class Classification

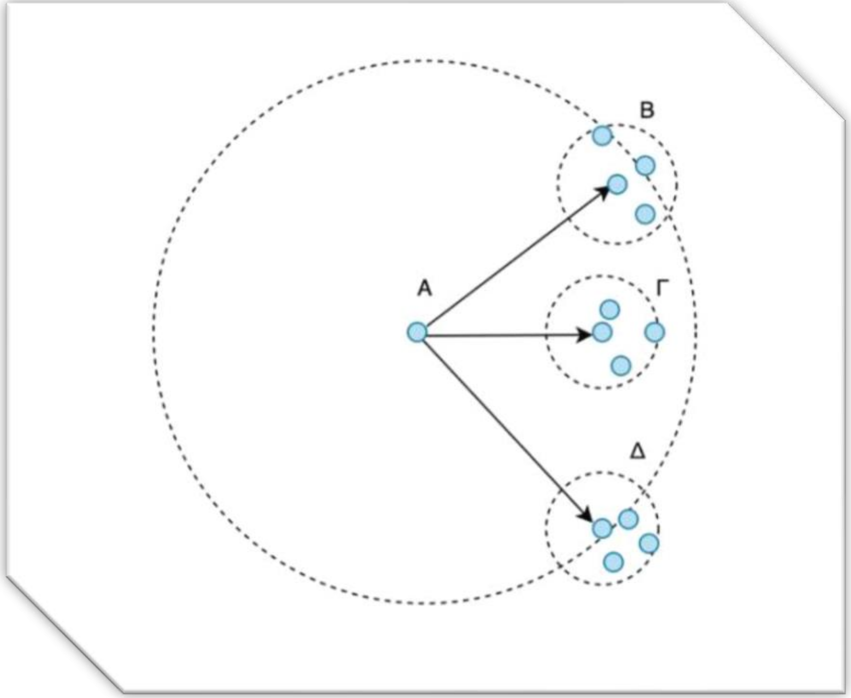
Detection of Extreme Samples (Outlier Detection):

- Unsupervised
- Detection of Areas of High Sample Density
- Data Denoising
- Isolation Forest, Elliptic Envelope, Local Outlier Factor

Detection of Unusual Samples (Novelty Detection):

- Semi-Supervised
- Delimitation of the Total Education Area
- Denoised Training Sets
- One Class Support Vector Machine

Local Outlier Factor (LOF)

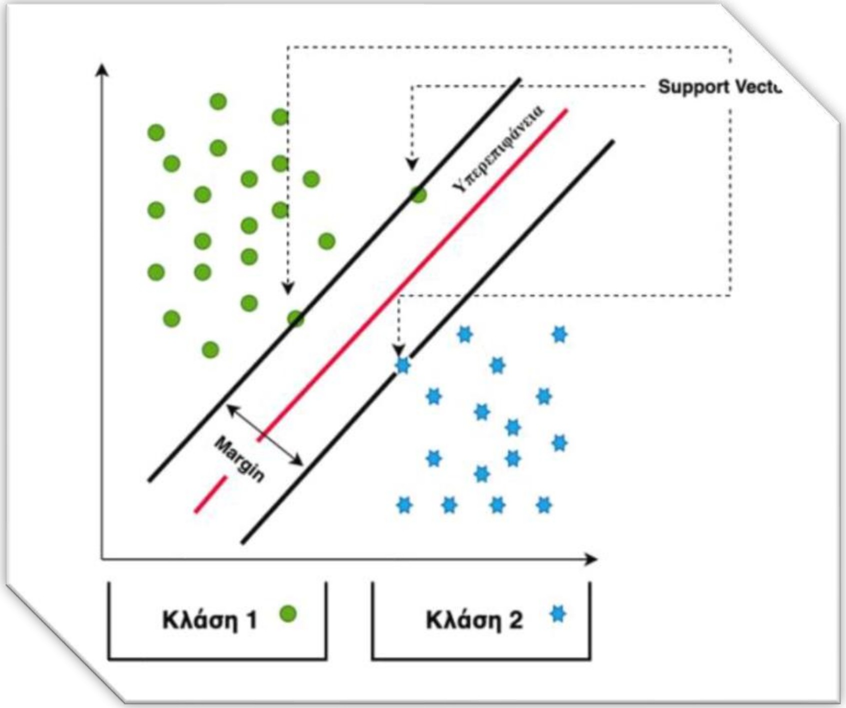


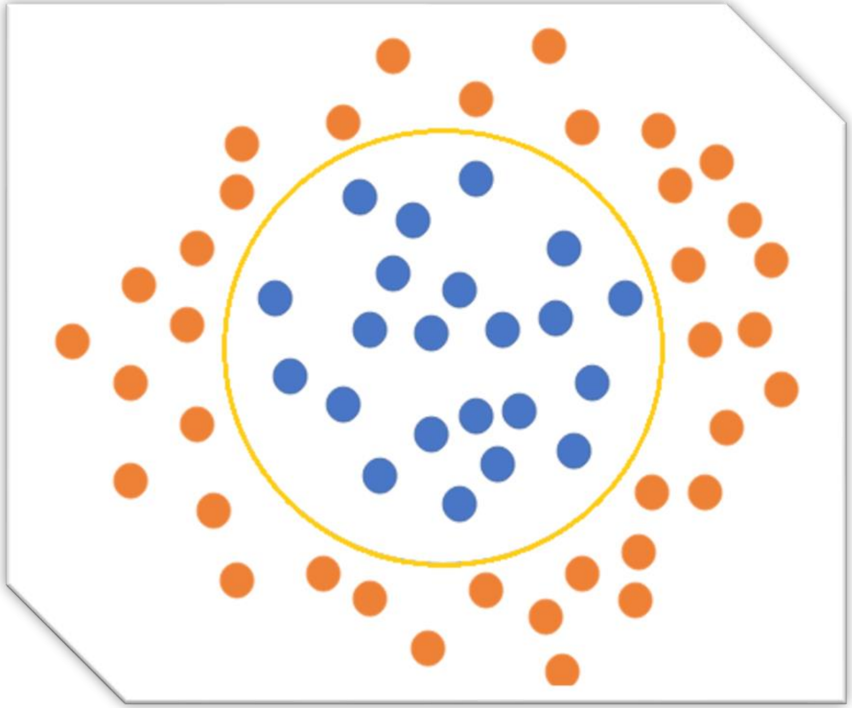
$$RD(Xi, Xj) = max(kDistance(Xj), Distance(Xi, Xj))$$

$$LDR_k(A) = \frac{1}{\sum_{Xj \in N_k(A)} \frac{RD(A, Xj)}{||\mathrm{N_k(A)}||}}$$

$$LOF_k(A) = \frac{\sum_{Xj \in N_k(A)} LRD_k(Xj)}{||\mathrm{N_k(A)}||} \times \frac{1}{LRD_k(A)}$$

One Class Support Vector Machine (OCSVM)





Evaluation Metrics

$$\text{False Rejection Rate} = \frac{FN}{TP+FN}$$

$$\text{False Acceptance Rate} = \frac{FP}{TN+FP}$$