issel
Intelligent Systems & Software Engineering Labgroup

ARISTOTLE UNIVERSITY OF THESSALONIKI

SOFTENG GROUP
Software Engineering Research

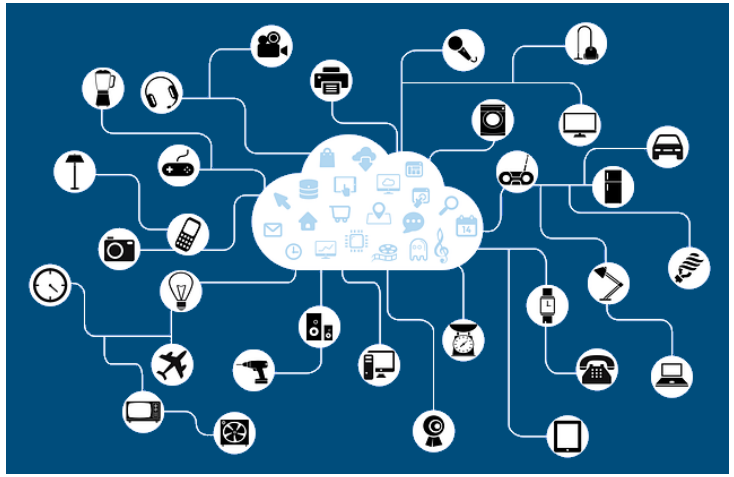
Softeng group
Intelligent Systems and Software Engineering Labgroup
Electrical and Computer Engineering Dept.
Aristotle University of Thessaloniki, Greece

Workshop: Data Analytics for IoT

Friday, October 23, 2020

Data Analytics for IoT

- Internet of Things is the network of **physical objects ("things")**, embedded with sensors, software, and other technologies, that **collect and exchange data**
- Data Analytics provides the means to **analyze data** in order to draw useful conclusions and **support decision-making**

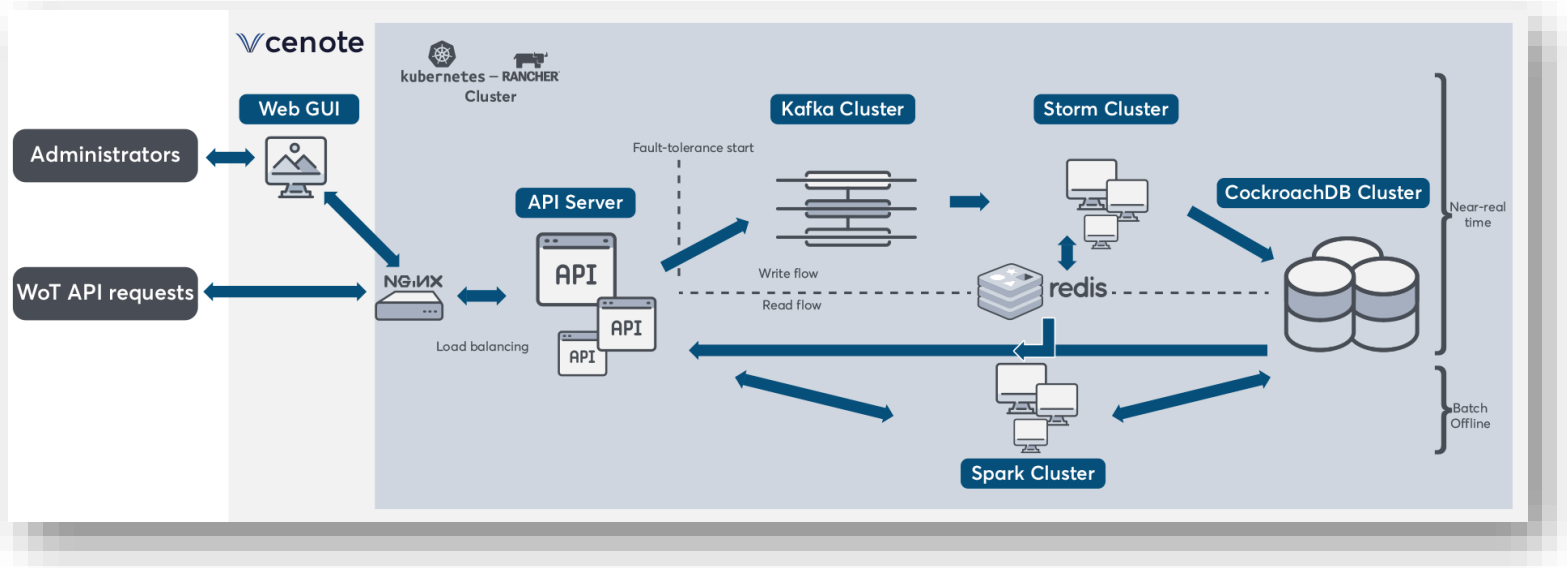


Programme

- Cenote: Big Data Management System and Applications
- BrainRun: Application on Behavioral Biometrics
- VITAL: Application on Precision Agriculture

Cenote: Big Data Management System and Applications

- Big Data Management System (BDMS) based on open source components
  - Deployed in a **distributed and scalable** manner
  - Analytics out-of-the-box in **event stream processing**
  - Support for **real-time analytics** as well as batch processing

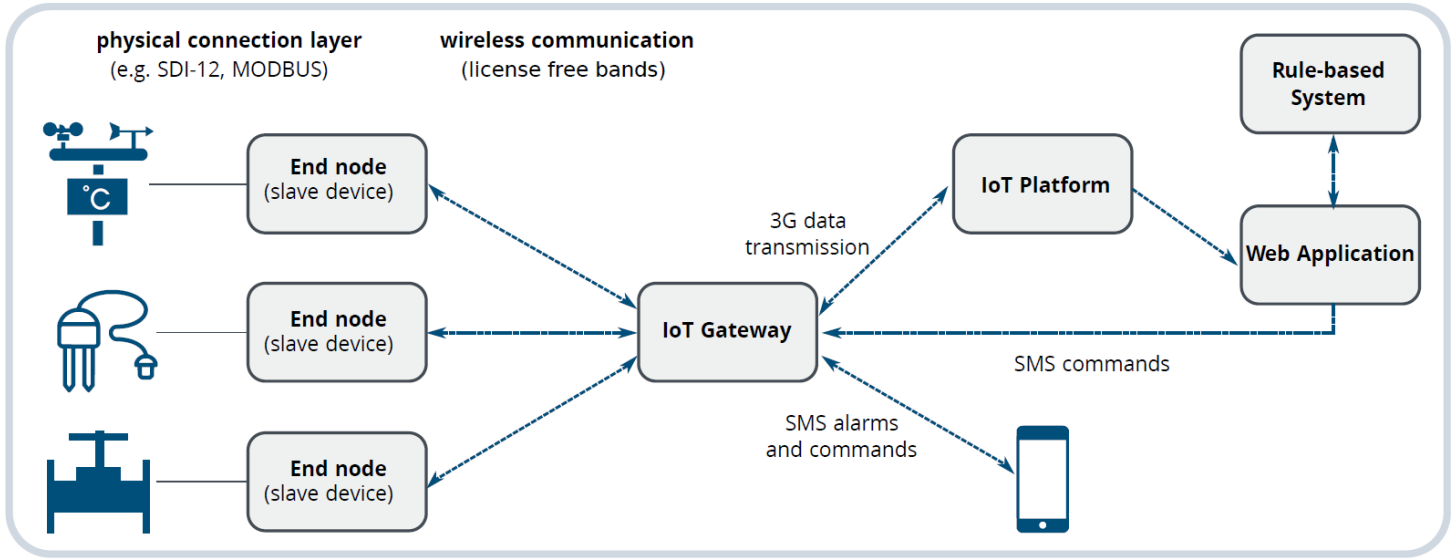


BrainRun: Application on Behavioral Biometrics

- Design and Development of a **Continuous Implicit Authentication** methodology based on the interaction of end-users with touch screens
- Data gathering:
  - A **brain training game** aiming at giving your cognitive skills a boost!
- Development of a machine-learning based authentication mechanism

VITAL: Application on Precision Agriculture

- Integrate AI and IoT for agriculture
  - Collect field data (**smart agriculture sensors**)
  - Process and store the data (**visualization & analysis**)
  - Decide and act (**data-driven decisions**)



Part A

Cenote: Big Data Management System and Applications

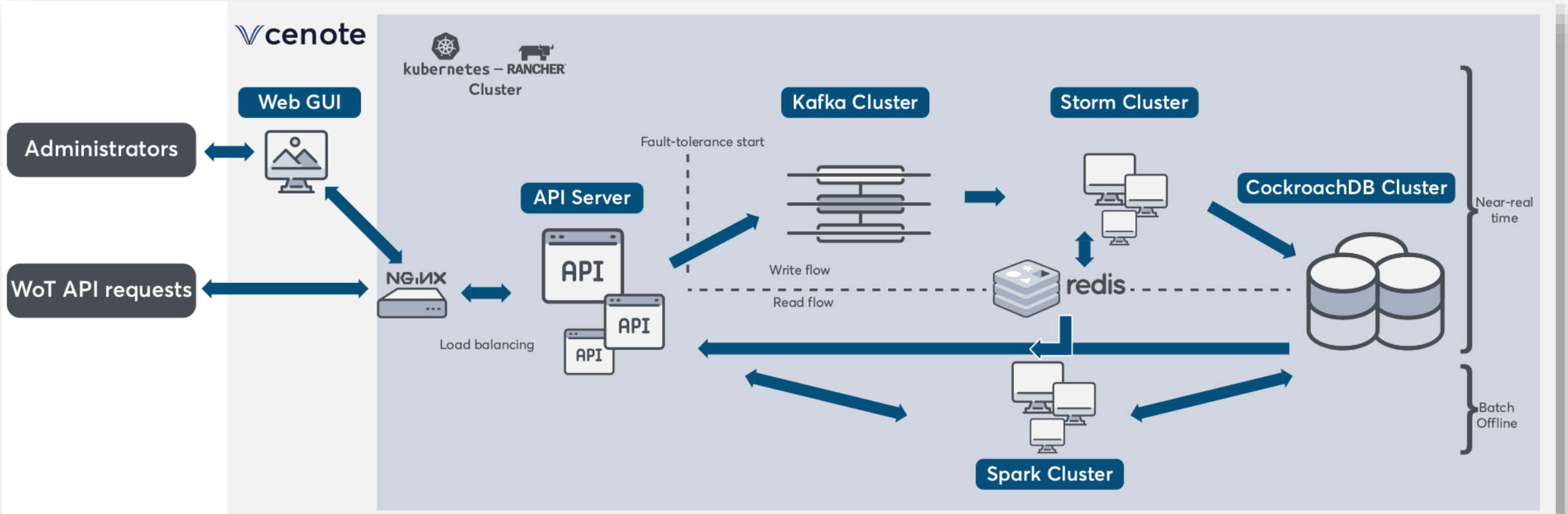
Cenote at a glance

- Open source system based on open source components.
- Big Data Management System (BDMS).

Key Features:

- General Purpose.
- Analytics out-of-the-box in event stream processing.
- Support for real-time analytics as well as batch processing.
- Deployed in a distributed and scalable manner

Cenote Architectural Overview



Data Modelling

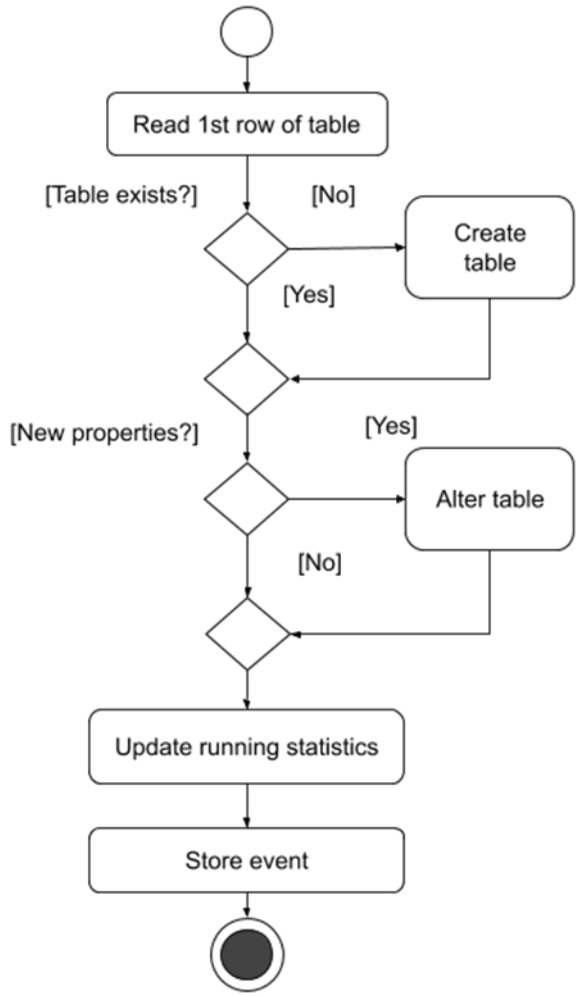
- Information organized in Projects and Collections
  - Access control using Read/Write/Master keys
- Automatic schema identification and update
- 2 types of data flow
  - Read Flow
  - Write Flow

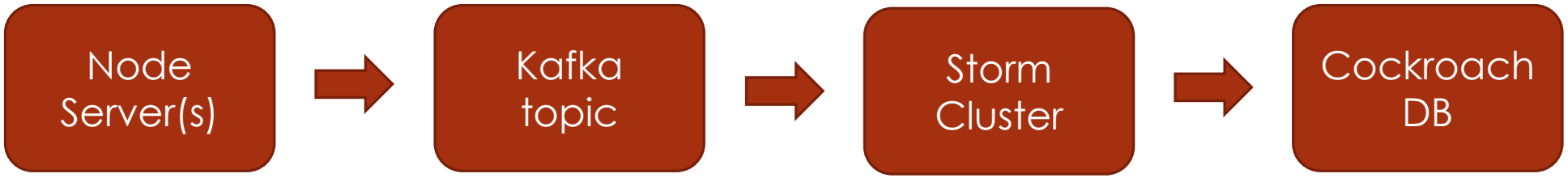
```
{
  "cenote": {
    "created_at": "2012-12-14T20:24:01",
    "timestamp": "2012-12-14T20:24:01",
    "id": "asd9fadifjaqw9asdfasdf939"
  },
  "device": {
    "id": "1234567890abcdef",
    "model": "H09 Beta",
    "temperature": 29.5
  }
}
```

Softeng group, Intelligent Systems and Software Engineering Labgroup
Electrical and Computer Engineering Dept., Aristotle University of Thessaloniki, Greece

Write Flow

- Incoming requests are served by 1 or more node servers running in multithreaded mode.
- Event-based handling on requests.
- Schema identification based on incoming data.
- Updates running statistics (mean and variance) for outlier detection per numeric property



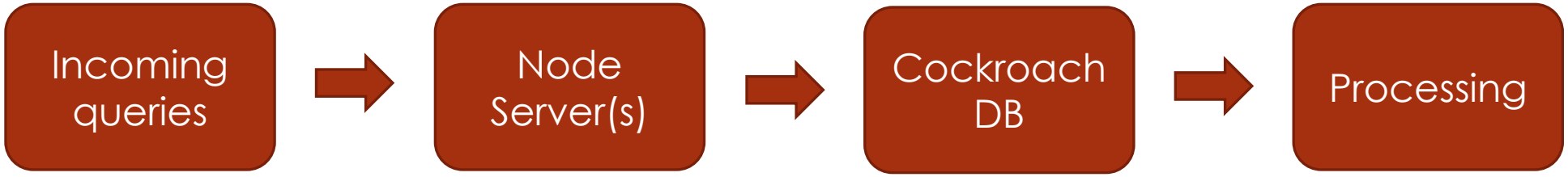


Read Flow

- Incoming API requests are translated into SQL Queries.
- Post-processing in node.js server for complex queries (e.g. count unique for 10% percentile)
- Time-based analytics for all properties. Supported Types:

  Average, Sum, Count Unique, Count, Maximum, Median, Minimum, Percentile, Select unique

- Group by capabilities



Evaluation

- With a 5 seconds critical path, production cluster was able to handle 1K requests/second end-to-end (from receiving to persisting the request).
- Architecture enables scalability in three levels based on individual needs:
  - Server
  - Workers
  - DB
- Benchmarking (using the current infrastructure) shoed that CENOTE can handle 25K requests/second.
- The online outlier detection algorithm was proven to have a small impact on the response times

Part B

BrainRun: Application on Behavioral Biometrics

BrainRun at a glance

- Continuous Implicit Authentication
- Touch Traces Modelling

Key Objectives:

- Dataset from real-life Application.
- Features identification.
- Construction of CIA Models.
- Kiosk vs Mobile devices

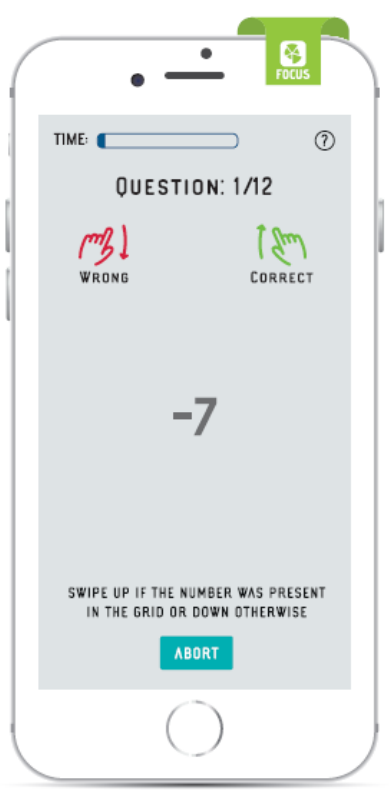
BRAIN RUN

BrainRun App (1/2)

- 5 different game types
  - Mathisis (horizontal freely-swiping pattern)
  - Focus (vertical freely-swiping pattern)
  - Memoria (force-tapping pattern)
  - Speedy (force-tapping pattern)
  - Reacton (force-tapping & force-swiping pattern)
- Solo mode including 160 games (5 stages with increasing difficulty)
- Tournament mode

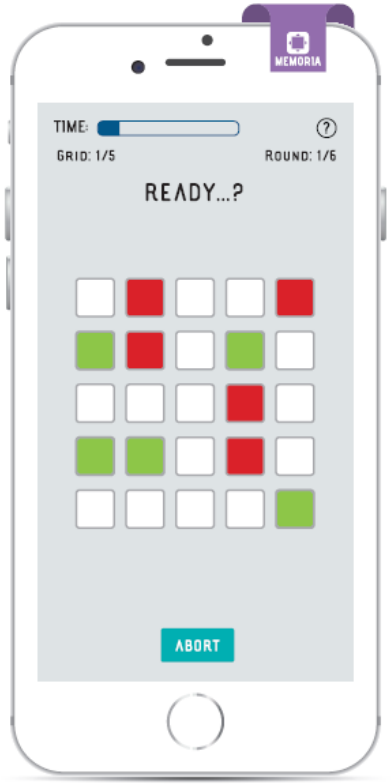
Softeng group, Intelligent Systems and Software Engineering Labgroup
Electrical and Computer Engineering Dept., Aristotle University of Thessaloniki, Greece

BrainRun App (1/2)

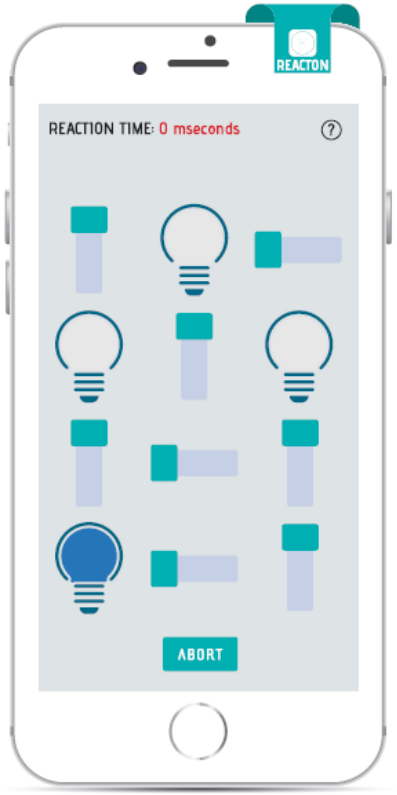


Focus

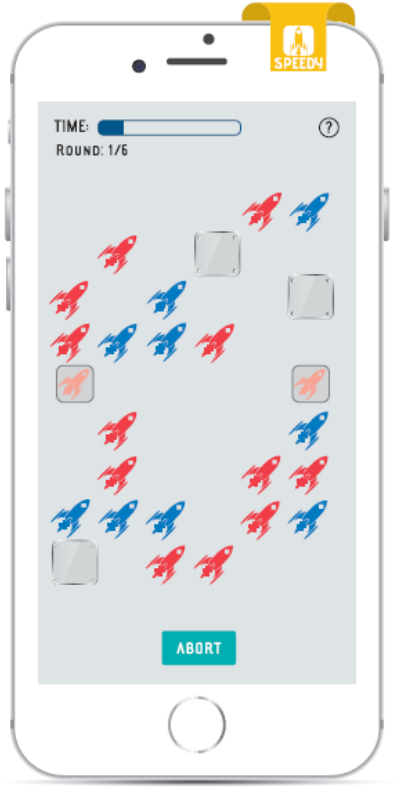
Mathisis



Memoria



Reacton



Speedy

Evaluation

- Gathered touch traces data for than 2,218 users and 2,418 devices
- More than 100K games played
- More than 3.5M swipes and taps
- Application of ensemble one-class classification methodology using Support Vector Machines
- Introduction of confidence for application in real-life application
- False Rejection Rate < 6%

Part C

VITAL: Application on Precision Agriculture

Objective of the VITAL project

- Field telemetry => wealth of information vis-à-vis:
  - Meteo data
  - Status of the soil
  - Irrigation needs

- IoT coupled with Big Data processing of Earth Observation data (i.e. Sentinel of Copernicus), *they can transform the modern agricultural domain*

Pilot application of VITAL to be examined today I

Pilot application of VITAL to be examined today II

Pilot application of VITAL to be examined today III

- Data recorded every 5 minutes
- 2 years worth of data
- Plot was split into 4 areas with different treatments
- End goal was to perform precision irrigation according to crop needs and soil status

- In our workshop we will
  - 1. Study techniques to examine the data, perform various types of aggregations, transform the data
  - 2. Build a dataset from the data and study the principles of machine learning

For more information about VITAL

- *"Precision Farming and the Internet of Things — The VITAL project"*
  - Final VITAL conference, Friday, October 23, 2020 (13:30–16:00 EEST)

- Visit https://www.vital-agro.gr/ for more!

Thank You!

Contact info

Themistoklis Diamantopoulos
thdiaman@issel.ee.auth.gr

Michail Papamichail
mpapamic@issel.ee.auth.gr

Thomas Karanikiotis
thomas.karanikiotis@issel.ee.auth.gr

Nikolaos Tsakiridis
tsakirin@ece.auth.gr