

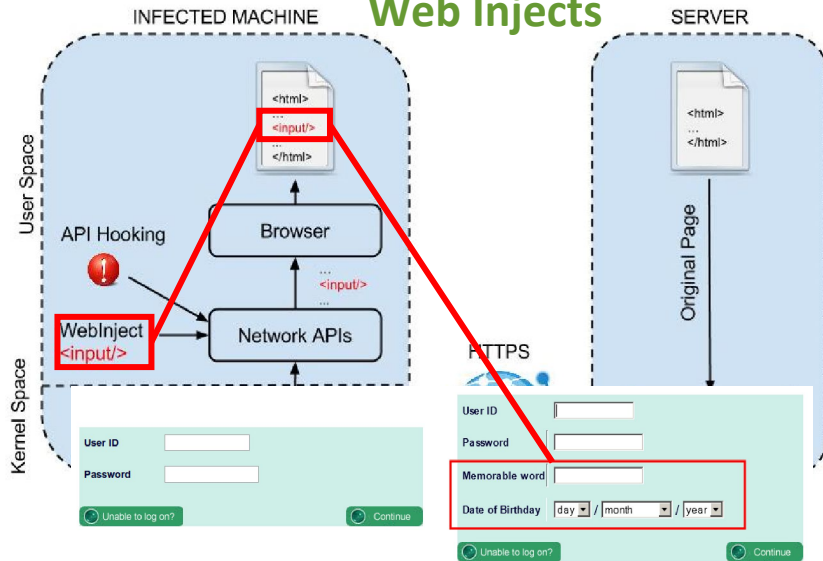


# Financial Fraud Detection



# Threats and Anatomy of a Fraud

## Web Injects



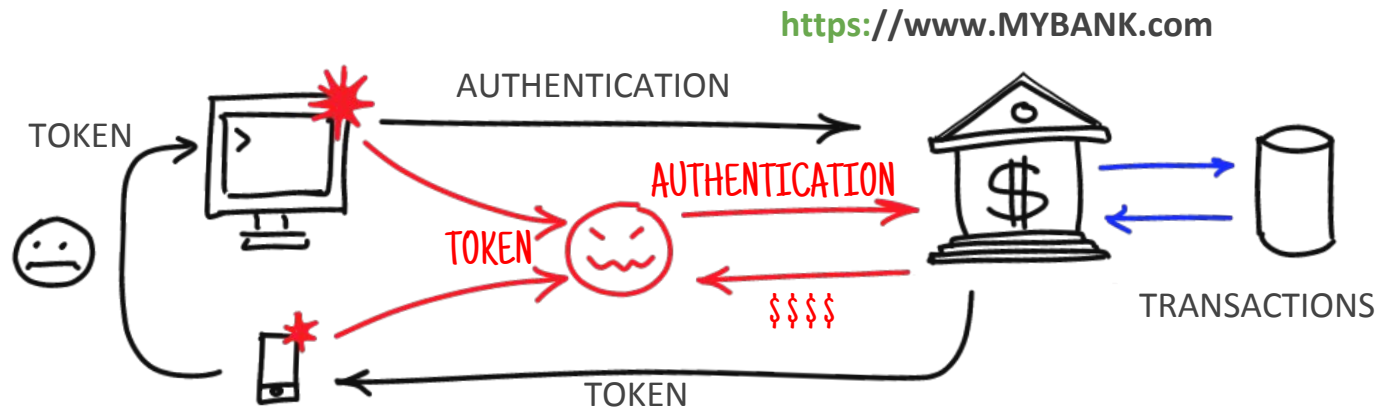
## Traditional threats:

- Phishing
- Credentials Database Theft

**Banking Trojans:** Malware that aims to perform online financial frauds.

## Man-In-The-Browser attack:

- steal credentials and private information
- Hijack browser session
- Infect Mobile Devices



# Internet Banking Frauds Challenges

Internet banking frauds are difficult to analyze and detect:

- **Fraudulent behavior** is dynamic and dispersed in large and highly imbalanced datasets with different customer's profile
- Scarcity of available informations and data
- Most of the **existing approaches**:
  - Black box
  - Based on Synthetic data
  - Not adaptive baseline profiling

# Goals

- Not focus on pure detection approach
- Support the analysis and the investigation of (novel) frauds and anomalies through readable model and results.
- Decision support system able to model user behaviour and its evolution

# Approach and System Description

Overview

User Profiles

Undertraining and Updating

# ATTRIBUTES

BANK TRANSFERS

\$\$\$ CC.IP IP IBAN NAZ\_IBAN D:H:M:S

PHONE RECHARGES

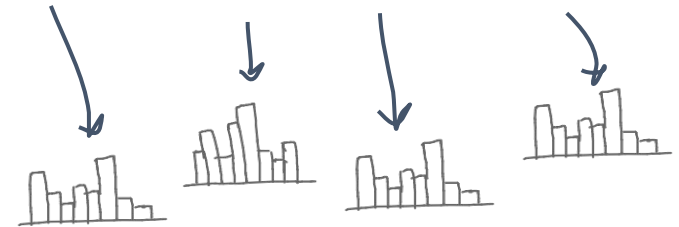
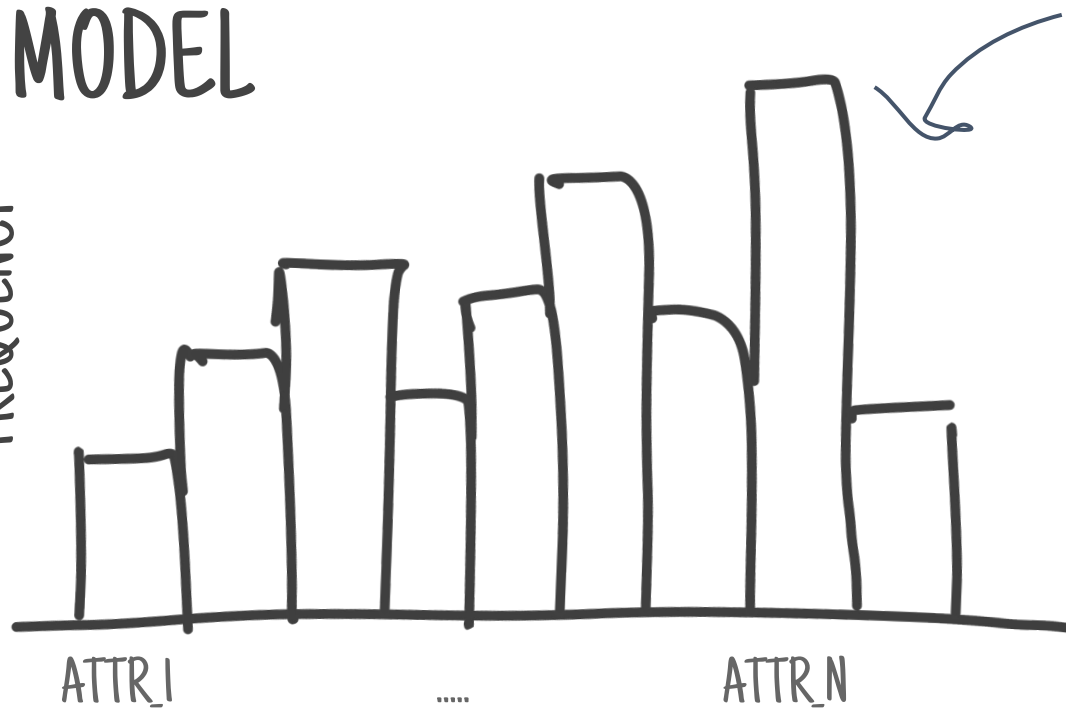
\$\$\$ CC.IP IP OP.TEL NUM.TEL D:H:M:S

PREPAID CARDS

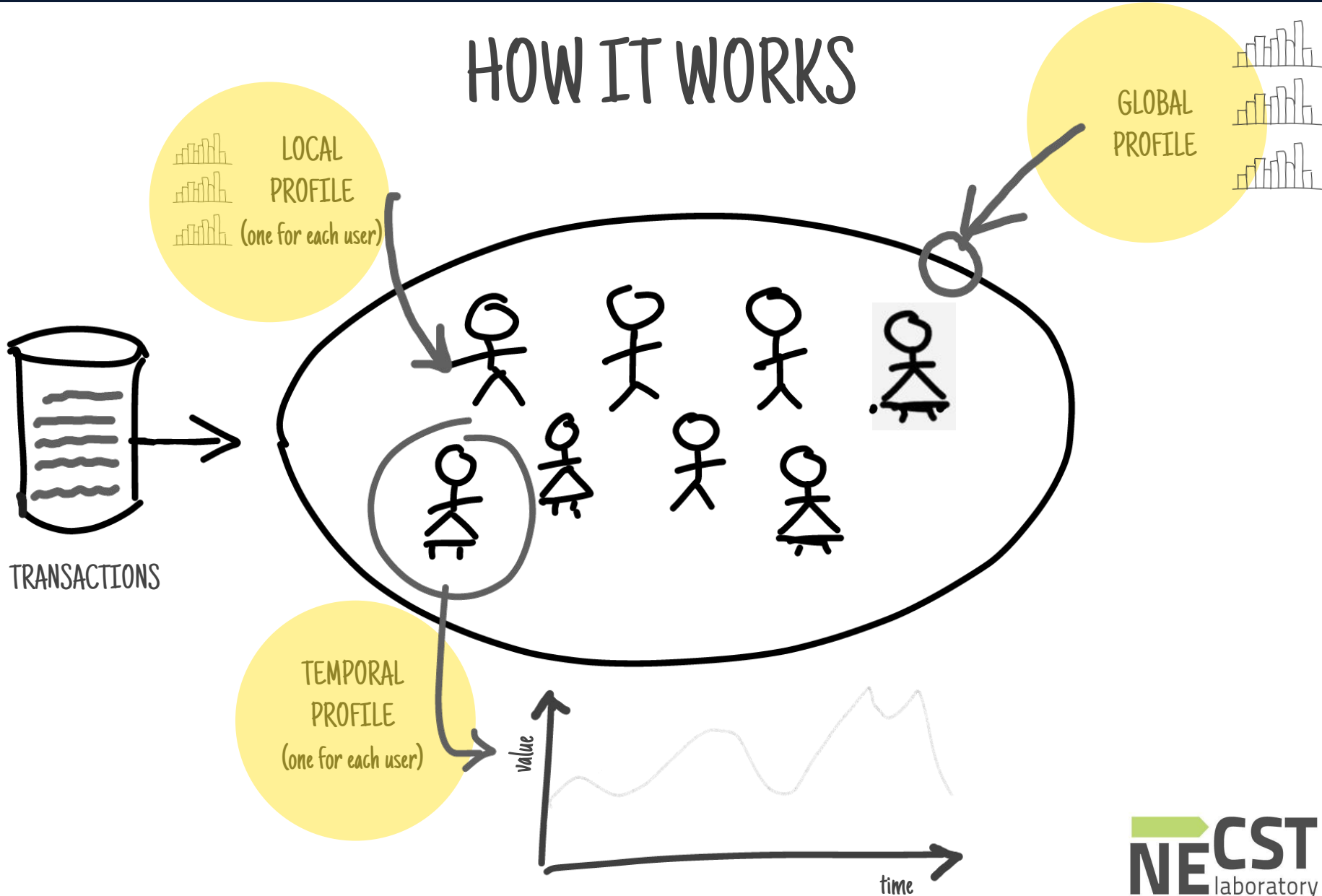
\$\$\$ CARD\_TYPE CARD.NUMBER CC.IP D:H:M:S

## MODEL

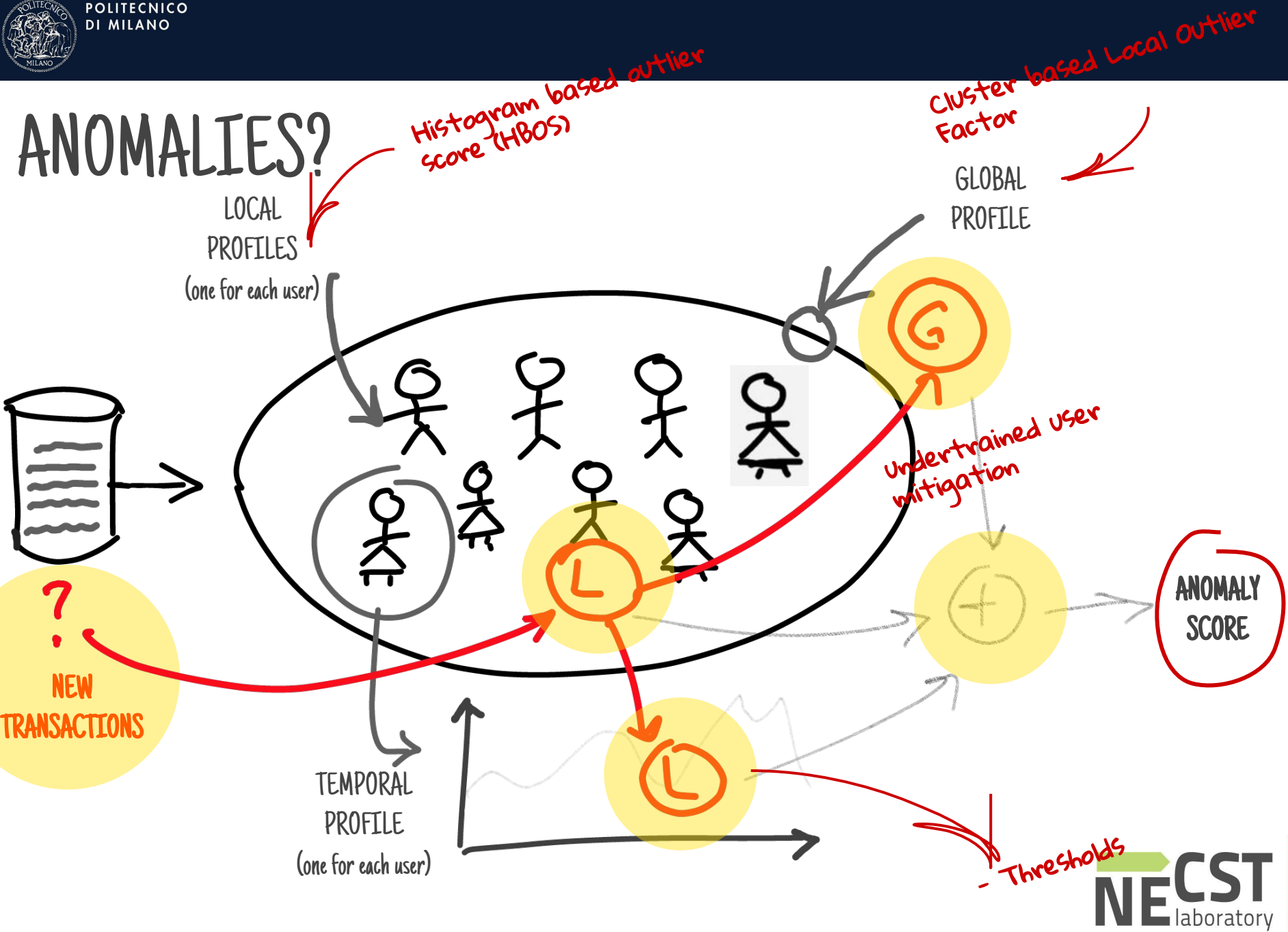
FREQUENCY



# HOW IT WORKS



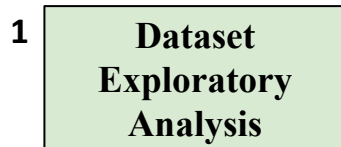
# ANOMALIES?



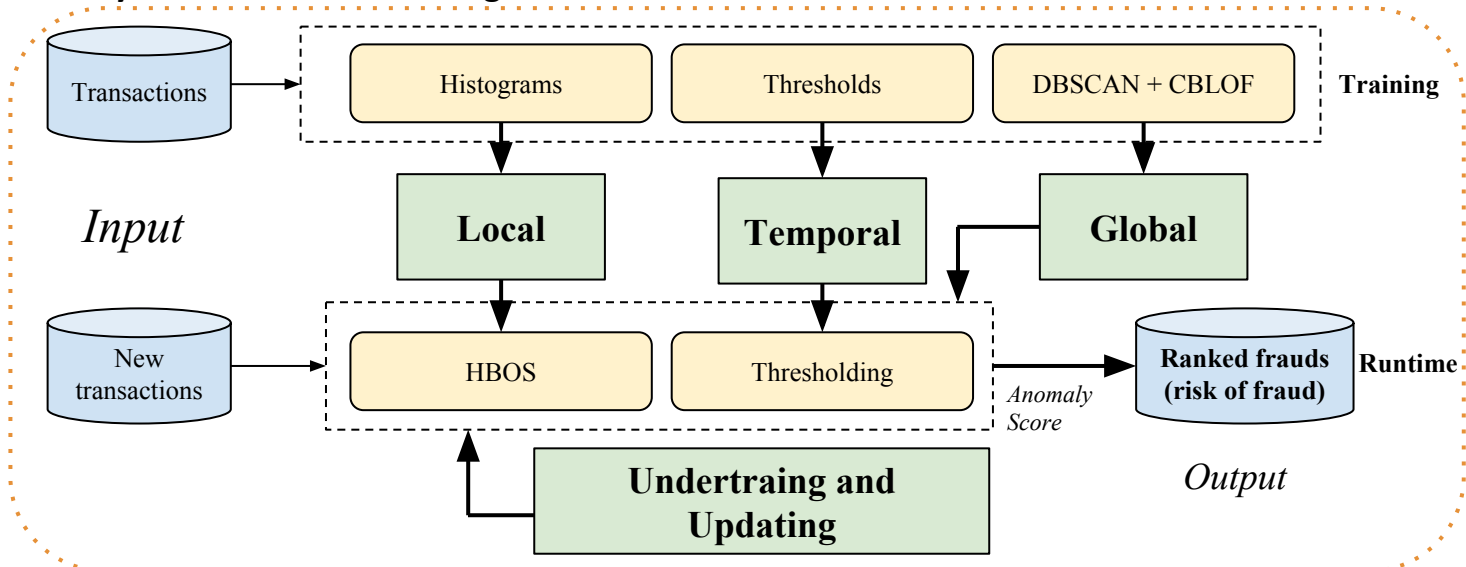


# BankSealer

**Decision-support and fraud-analysis system** able to effectively rank frauds and anomalies

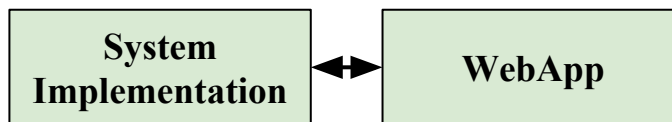


## 2 System Architecture Design



## 3 Implementation

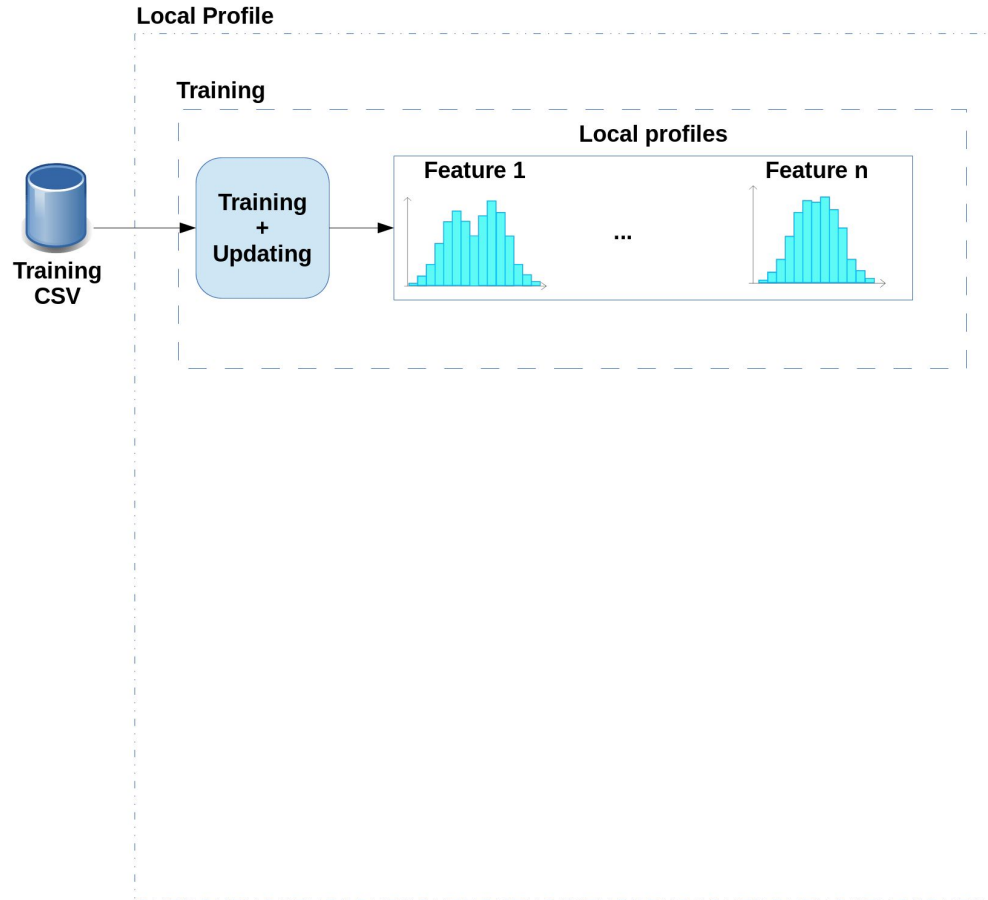
4 Testing and Evaluation



## 5 Deployment

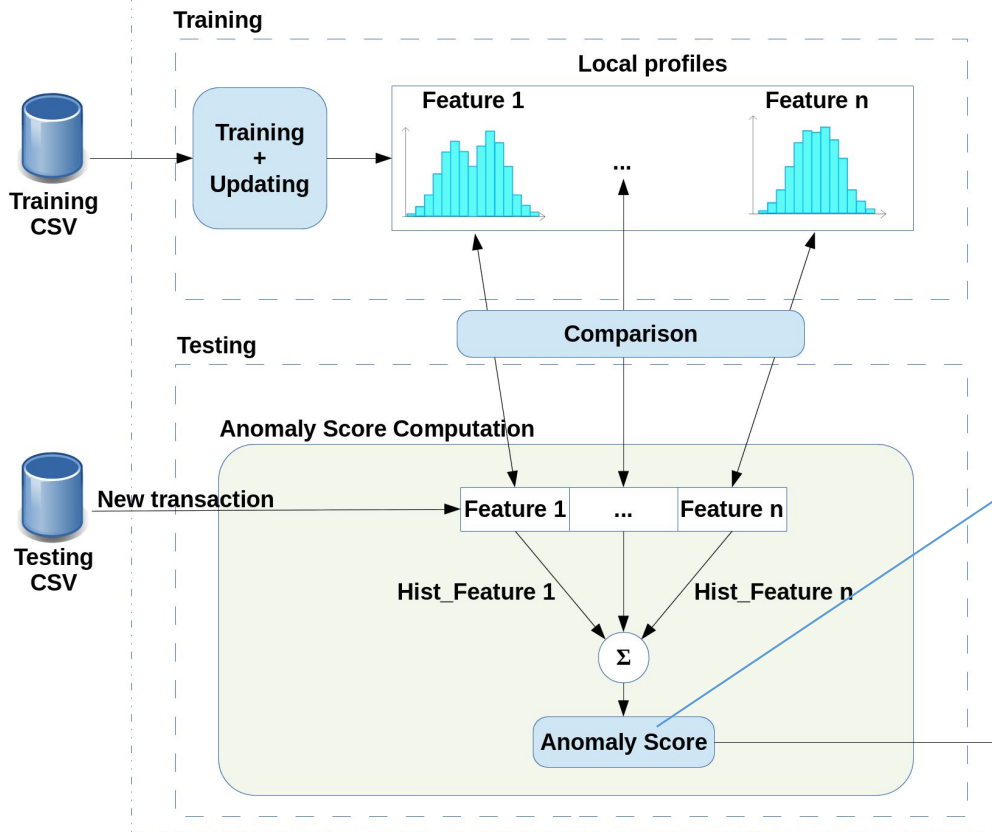


# Local Profile



# Local Profile

## Local Profile



## Histogram-Based Outlier Score

$$HBOS(t) = \sum_{0 < i \leq d} w_i * \log \frac{1}{f(t_i)}; \quad \sum_{0 < i \leq d} w_i = 1$$

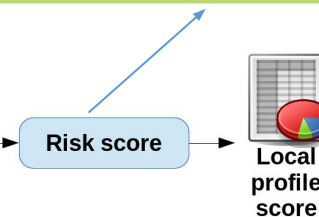
## Feature Normalization

$$f(t_i) = \frac{hist_i(t_i)}{\max_{j \in feature_i} hist_i(j)}$$

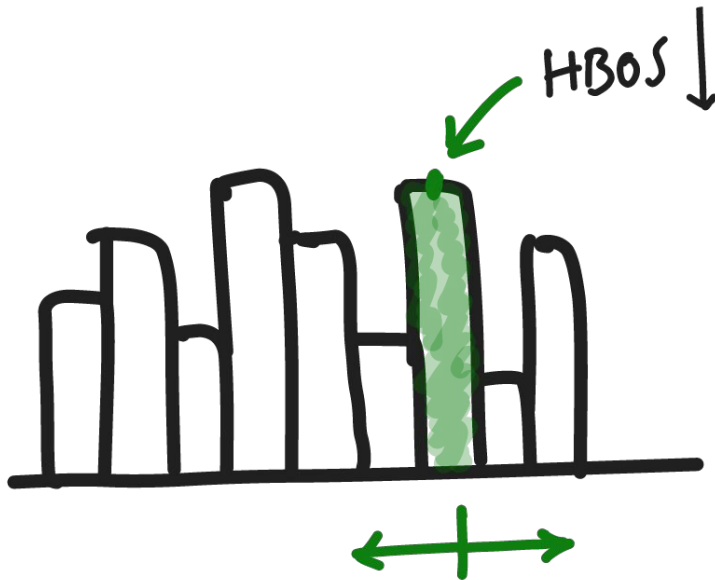
## Rare Values

$$\beta = \frac{k}{1 - f}$$

$$RISK(t) = t_{transaction\_amount}^{w_{amount}} * HBOS(t)^{w_{HBOS}}$$



# HBOS = Histogram Based Outlier Score



Each feature is weighted



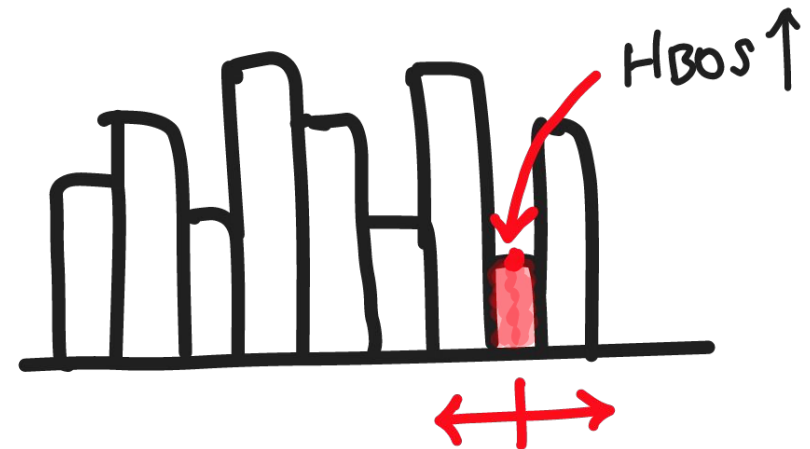
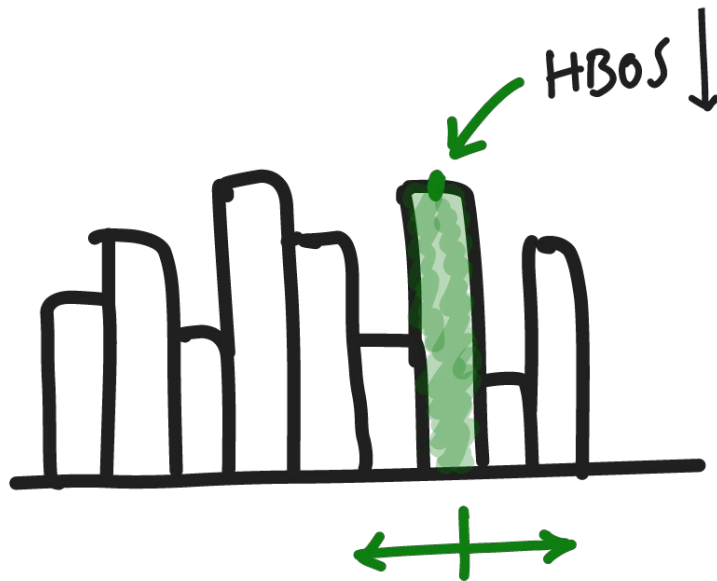
Prioritizes relevant features

Discount factor



Forget old data & Model concept drifting

# HBOS = Histogram Based Outlier Score



Each feature is weighted



Prioritizes relevant features

Discount factor



Forget old data & Model concept drifting

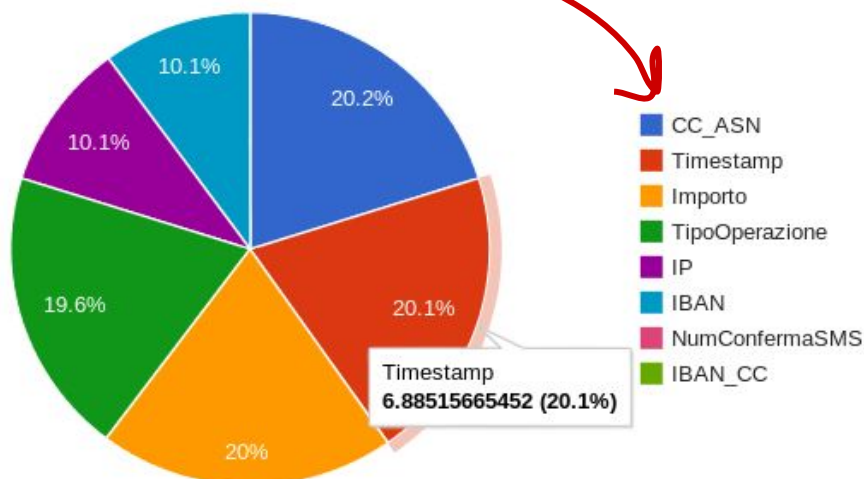
## HBOS visualization

index	IP	Timestamp	TipoOperazione	Importo	UserID	IBAN	Numero Conferma SMS	IBAN_CC	CC ASN	HBOS locale	Undertrained	New User
1	92580 3d64e9f4a188aa034659d1409f90456a	08/feb/2013 21:06:30	Giroconto	20000	dcfc15d4d65e05ebafde6ac9383062aa	e6c2a617f55090de28a24c67dfbedf40	✓	IT	IT,8612	29.402365073	✓	X
2	91133_99ca402ce2299ecd72e2ebe269b5d35f	06/feb/2013	Bonifici per detrazione	9900	4a4ee6e2ac1b17e20958ad7a8221c1b4	11e6c83f02065b037507b0b451144b	✓	IT	IT,3269	27.0032775111	✓	X
3	3c9dad3b2601c929e2cece7	✓	IT	IT,30722	24.5660880611	✓	X	5475914	X	X		
4										5765096	X	X
5	101355 86419f50fbda2742c1dba87cd3429476	28/feb/2013 17:46:46	Giroconto	12000	49zee7251c425c36c82cd3241c563f79	484fe271f1804b3e4291a537bb65279a	✓	IT	IT,44957	25.0995700682	✓	X
6	92502_dd5d85da0532104875e18e4e32bc152c	08/feb/2013 16:11:24	Bonifici per detrazione fiscale	3863.29	ac6989c1fae1085def1308e532082cb7	73b5047423c9dad3b2601c929e2cece7	✓	IT	IT,30722	24.5660880611	✓	X
7	99074_cd002daddde353900cc24e4ffc3b235c	21/feb/2013 18:40:02	Bonifici per detrazione fiscale	5643	a7b7a36b2769a1be86d1a544b67007a9	2626bfb3376dababb639201c9b8ff67	✓	IT	IT,21056	24.4493640116	✓	X
8	99827_0bdda3afaf28f049483d89d53f021c11	25/feb/2013 09:36:12	Bonifici Italia e SEPA	31000	be2b61118c081429cfbbc0c3d948743b	831687c224f781f106604f984e14f414	✓	IT	IT,12874	24.4175445119	✓	X
9	89586_2aeddb8850ae946914285eb3bcd28d55	04/feb/2013 16:42:53	Giroconto	10000	41efb45d969e9511b7df6504840cc572	40c200429a2c2a4c7268b3300681e5e3	✓	IT	IT,12874	23.6879134642	✓	X
10	101627_70c765c7265d92f96a05d91eebb4eb64	28/feb/2013 19:04:37	Bonifici Italia e SEPA	6529.6	8b7ed02e24a297a7ad7b91d28a5b35e1	3ada9624925ed42838bd4b8fab9eae81	X	IT	IT,50809	23.6370584204	✓	X
11	98401_d00d1939b4f71eaa199a57fff9cf0c19	20/feb/2013 14:59:10	Bonifici Italia e SEPA	50000	9bc3d0e6065284891a42ce6f9d828c38	65ecb9d1169b23049ec018d31c27af0a	X	IT	IT,3269	23.5882551789	✓	X
12	95342_2c2c6f325c547ee1fb0efc01475bc7d6	14/feb/2013 09:17:53	Bonifici Italia e SEPA	50000	f2a7341750c1cc6dc8bea45185a7fe26	60414014d030aa24b4cef90c32fac61f	✓	PT	IT,16232	23.5439265229	✓	✓
13	92842_ba3664bb7ebbf9e8bf4ac0664d65e239	10/feb/2013 19:21:11	Bonifici Italia e SEPA	20000	2a17ed71d9e2c82f39e174e424bf7eb9	e9987193889c72a6dcb94bbd47e35699	X	IT	IT,3269	23.5233646465	✓	X
14	92551_d7ab9d7839eb60ff6e06c496e1c848a	08/feb/2013 19:23:46	Bonifici Italia e SEPA	25266.8	435b8226966d2fb40d52baf6aaa8a93	92a91621f34b668401e8c26050f4e0c6	X	IT	IT,3269	23.4602697196	✓	X
15	97221_6da1465327246224216c1c929c339c6f	18/feb/2013 15:09:11	Bonifici Italia e SEPA	50000	f2a7341750c1cc6dc8bea45185a7fe26	60414014d030aa24b4cef90c32fac61f	✓	PT	IT,16232	23.2738702047	✓	✓

## HBOS visualization

## Local profile details

User feature contributions



motivations

attributes

Timestamp  
6.88515665452 (20.1%)

Close

Number of transactions: 3  
Final HBOS: 25.8683920672  
[Show user profile](#)

# Global Profile

Two phases:

## 1. Clustering

- a. **Algorithm:** incremental DBSCAN with decreasing epsilon
- b. **Distance Metrics:** Mahalanobis

## 2. Unweighted CLUSTER BASED LOCAL OUTLIER FACTOR: anomaly score based on the distinction between:

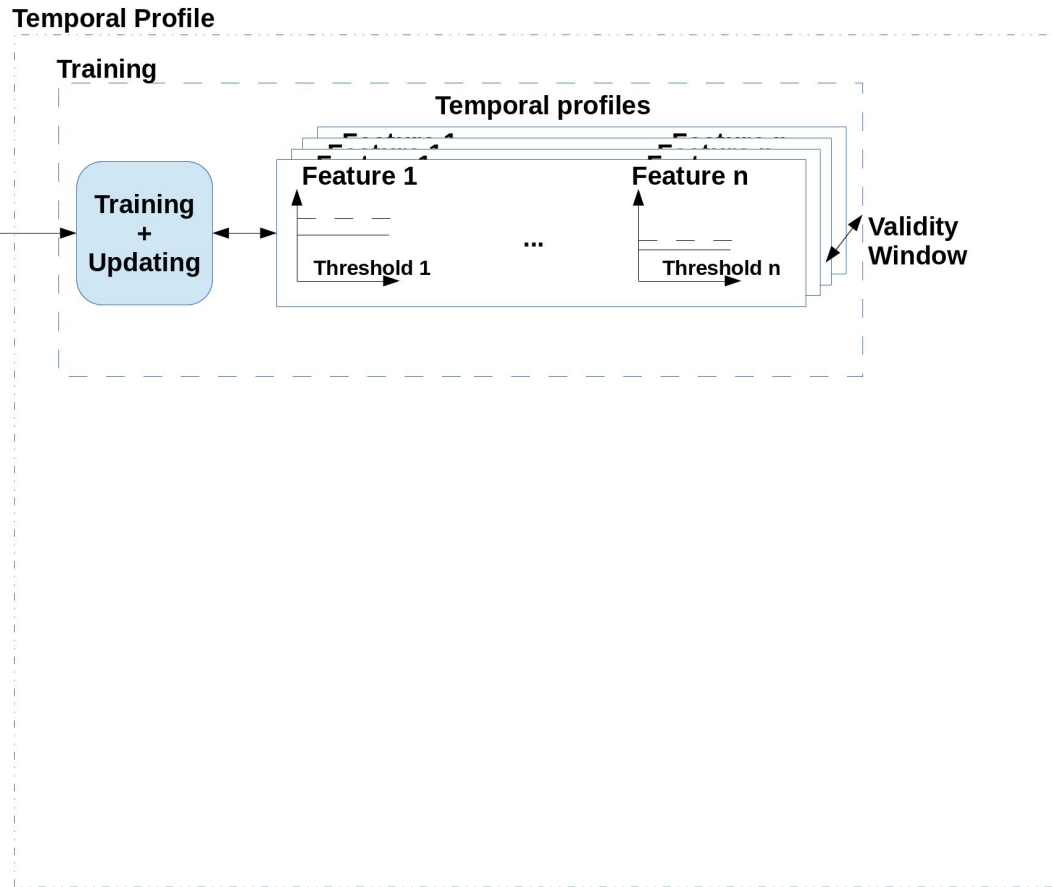
- A. LARGE Clusters LC
- B. SMALL Clusters SC



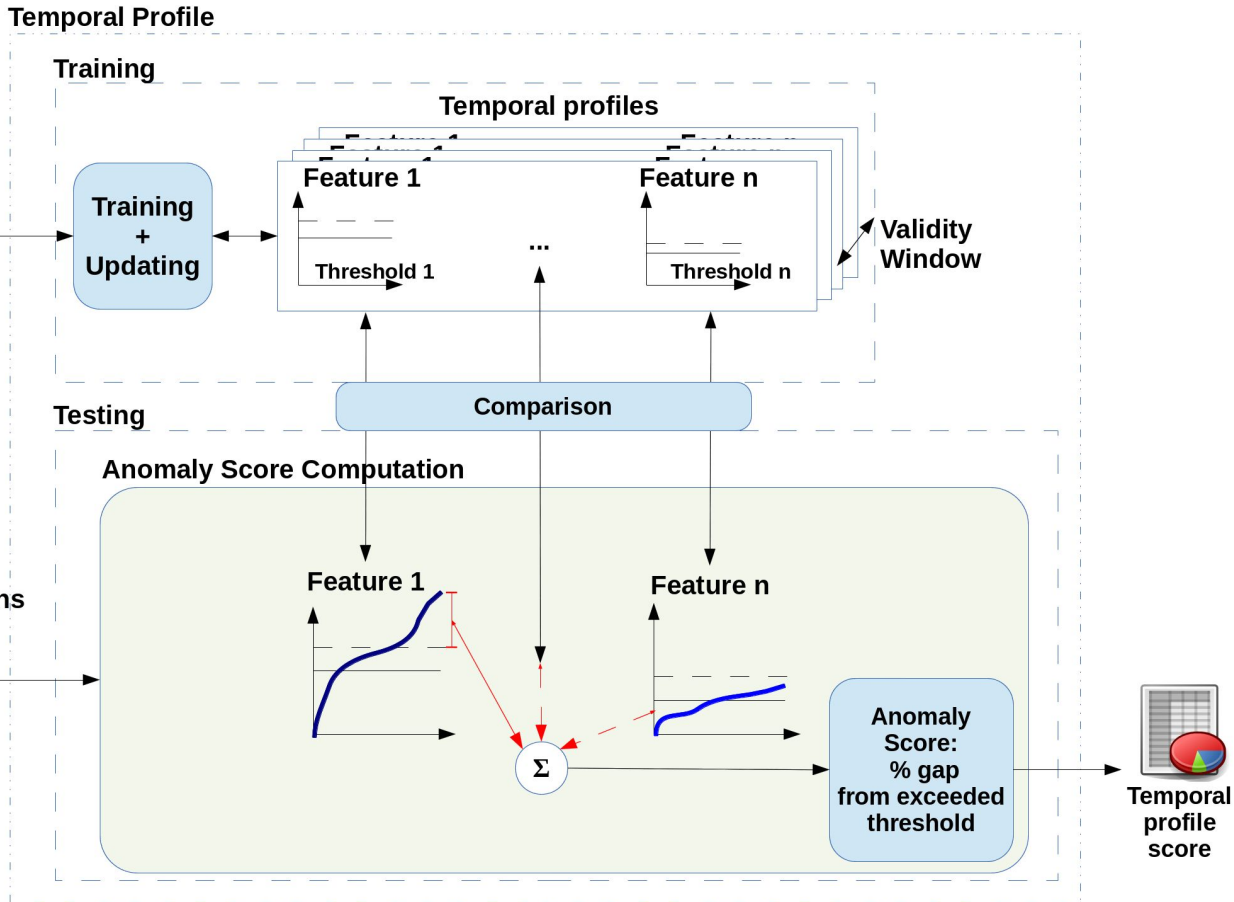
**CBLOF** = Min distance of a point from the centroid of the nearest LC



# Temporal Profile

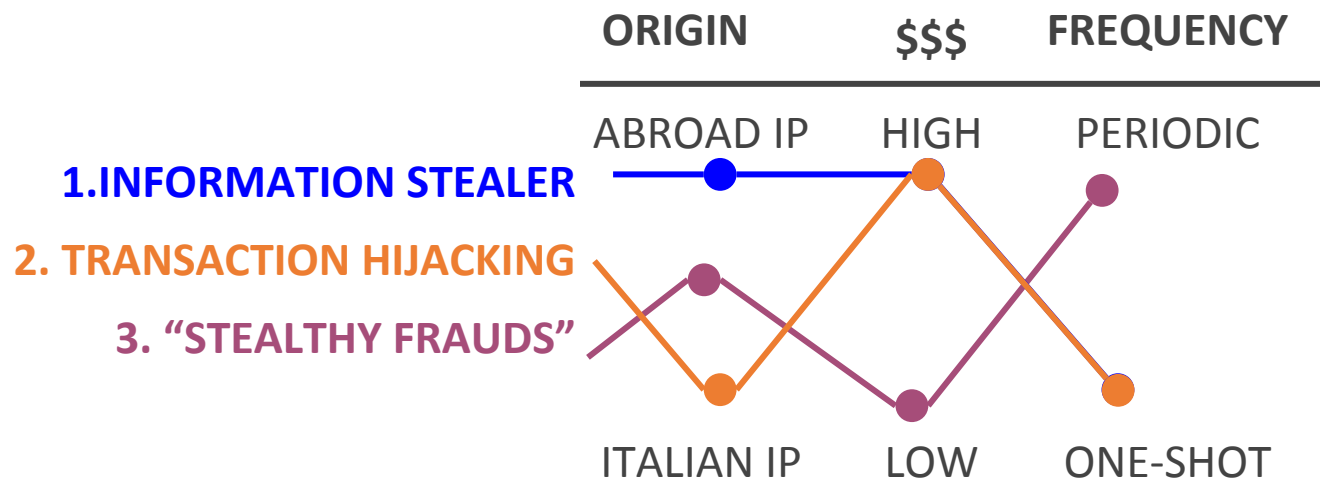


# Temporal Profile



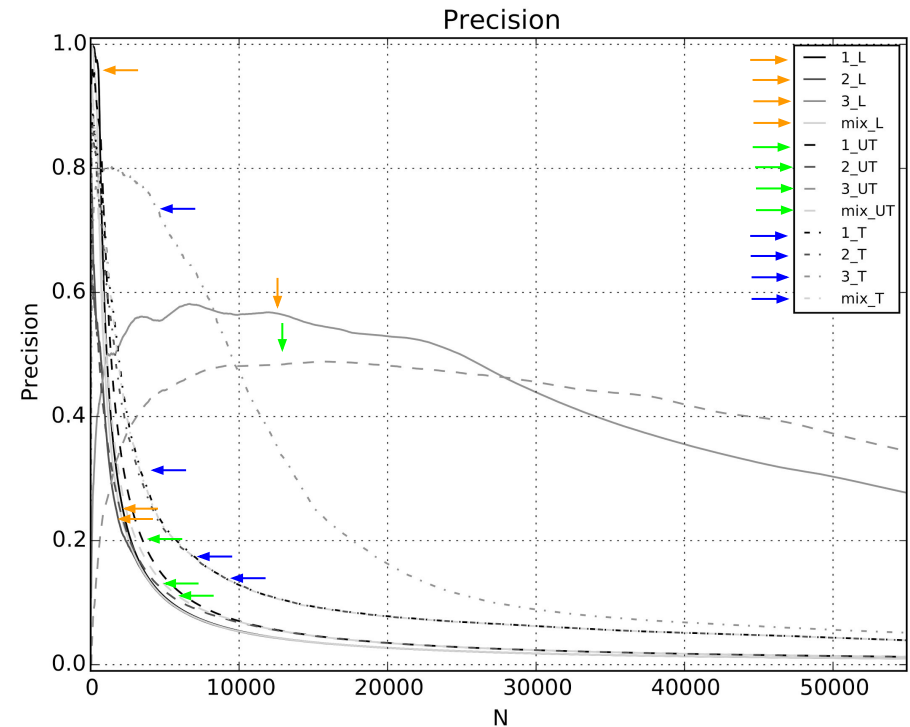
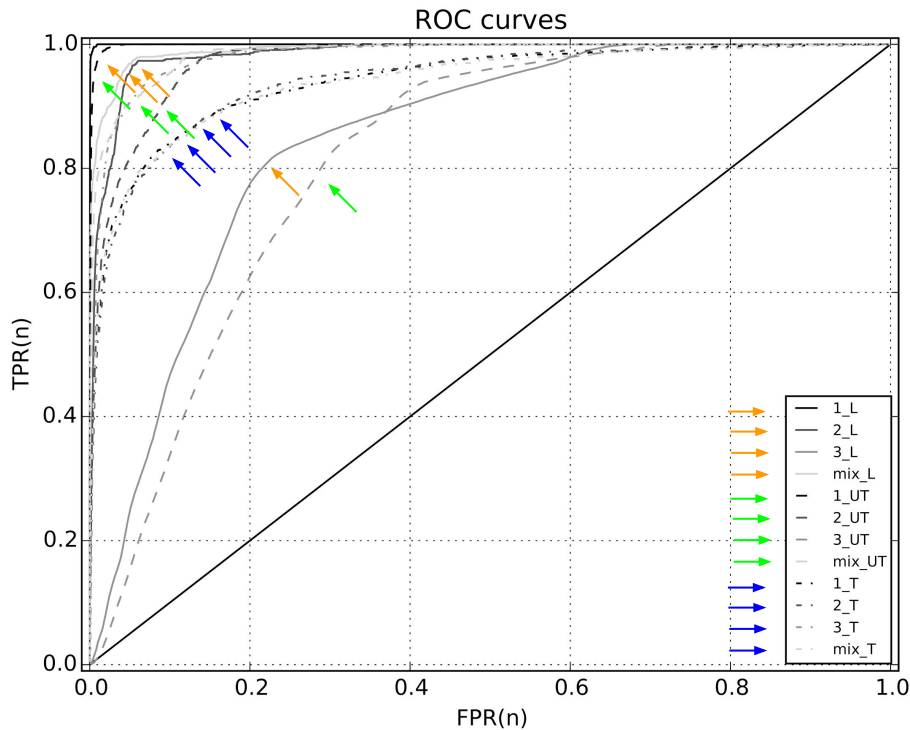
# Evaluation of BankSealer: Performance

Generate synthetic frauds based on scenarios built with the collaboration of bank experts that replicate the typical real attacks performed against online banking users



Inject **n fraudulent transactions (or users)** in the testing dataset and analyze the top **n transactions (or users)** in the ranking: ROC Curve - Precisions

# Evaluation of BankSealer: Performance



MICHELE CARMINATI, R. Caron, I. Epifani, F. Maggi, S. Zanero.

"BankSealer: A decision support system for online banking fraud analysis and investigation" Computers & Security, Volume 53, September 2015, Pages 175-186, ISSN 0167-4048,

<http://dx.doi.org/10.1016/j.cose.2015.04.002>

# FraudBuster: Temporal Analysis 2.0

## The Problem

Improve detection of frauds that exploits the repetition of legitimate-looking transactions

## Objective

Find a model able to describe the user's transactions in term of their periodicity and detect frauds as “deviations” from the learnt temporal model

## Proposed Solution

- Study **transactions distribution in the time domain** and classify user periodicity
  - Auto-correlation



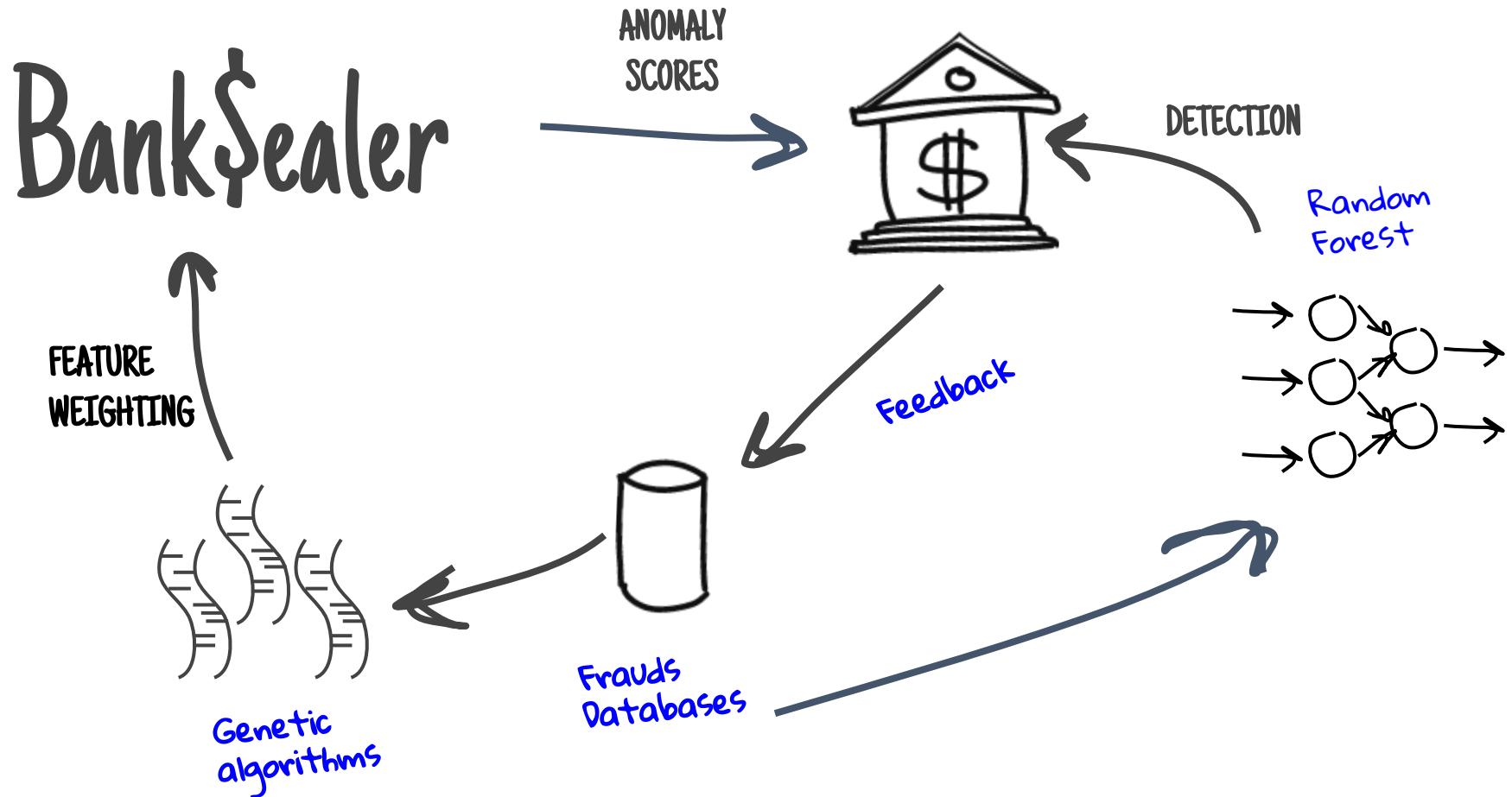
Almost 60% of user shows a monthly periodic behaviour

- **Detection**
  - Improved thresholds based on mean and variance
  - Histogram based distribution on monthly basis
  - *Dynamic Time Warping* on Transaction Time series



+ 30% detection rate improvement

# Supervised Analysis



# Thanks!

For further information:

[stefano.zanero@polimi.it](mailto:stefano.zanero@polimi.it)

[michele@banksealer.com](http://michele@banksealer.com)