POLITECNICO
MILANO 1863
Dipartimento di Elettronica
Informazione e Bioingegneria

POLITECNICO DI MILANO

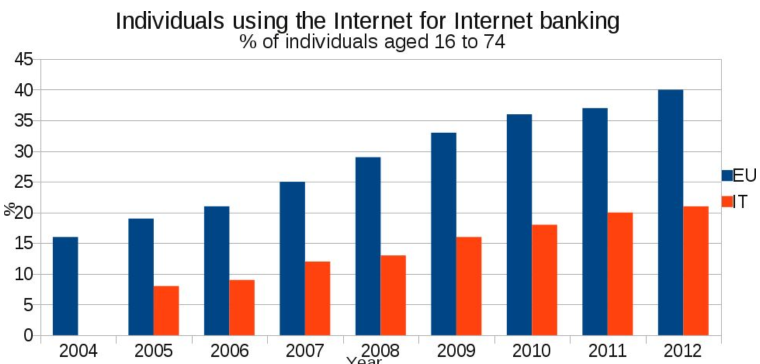
BankSealer: Fast and Transparent Online Banking Fraud Detection and Investigation

Federico Maggi - federico.maggi@polimi.it

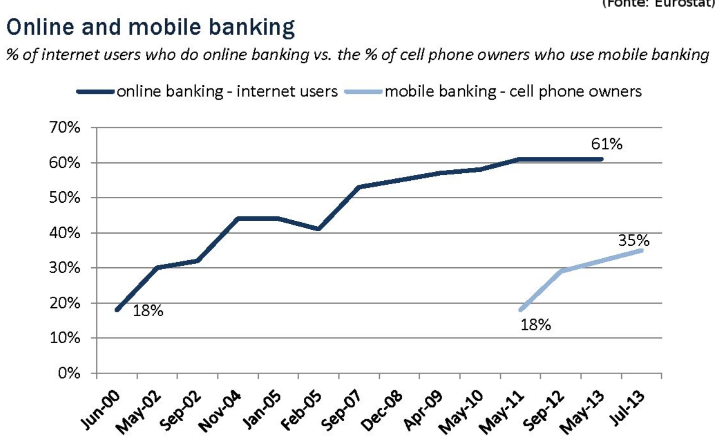
Joint work with: Michele Carminati, Stefano Zanero, Ilenia Epifani

NECST laboratory

Internet Banking



(Fonte: Eurostat)

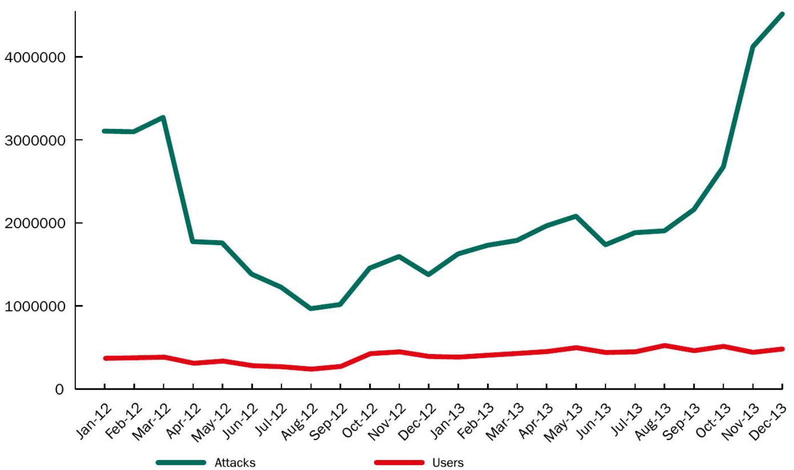


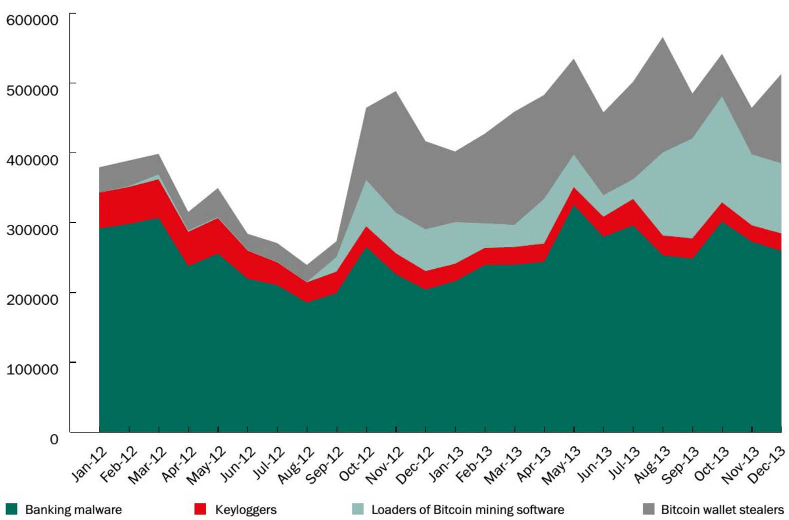
Source: Pew Research Center's Internet & American Life Tracking and Omnibus Surveys, 2000-2013. Margin of error for results based on internet users is +/- 2.5 percentage points and +/- 3.8 percentage points for results based on cell phone owners.

Growth of Internet banking services

Internet Banking

Financial malware: attacks and attacked users in 2012-2013



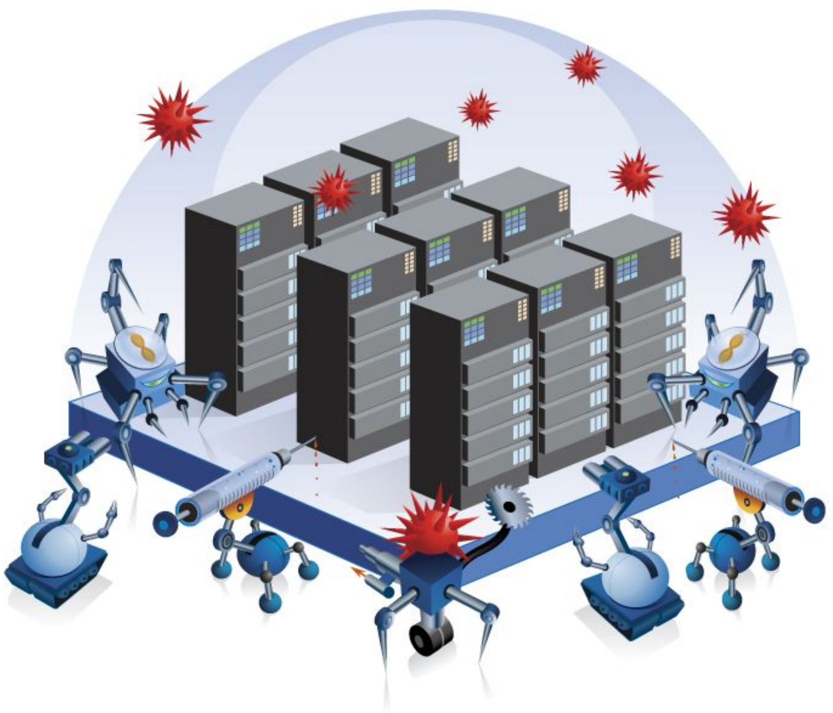


Growth of Internet banking services

⇩

Increase of online banking frauds and financial malware attacks

Internet Banking

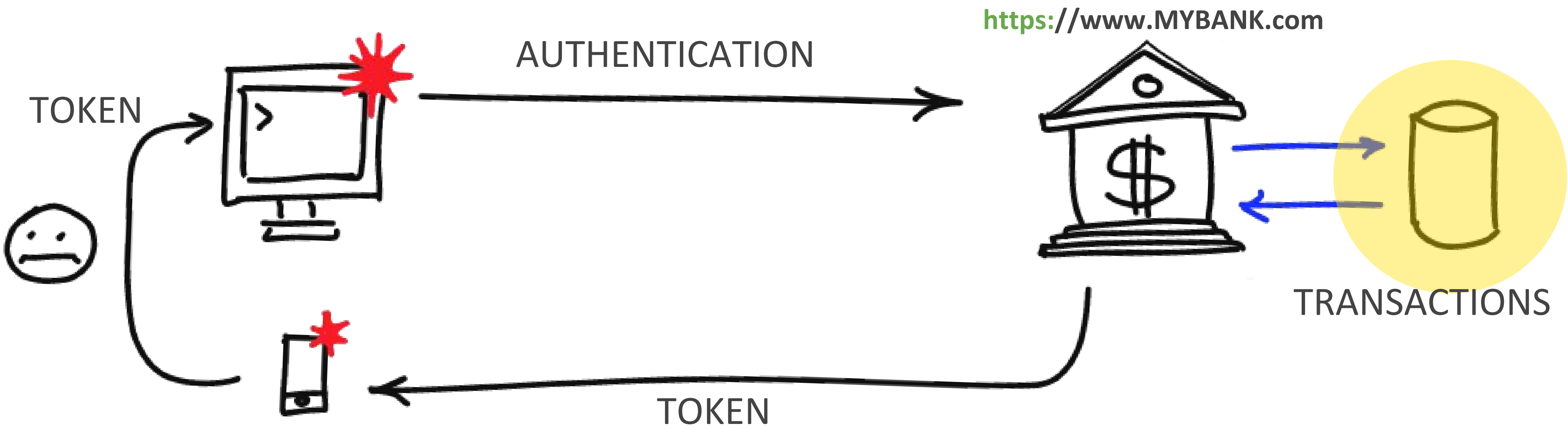


Growth of Internet banking services

Increase of online banking frauds and financial malware attacks

Need to create up-to-date <u>defense infrastructures</u>

Anatomy of a Fraud



Main threats

Traditional threats:

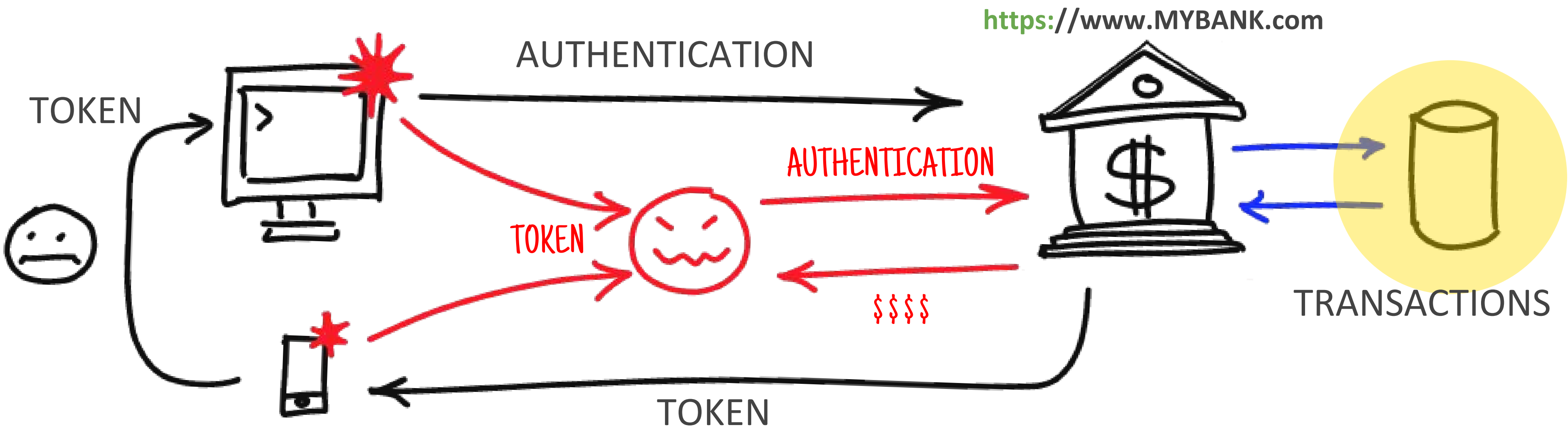
- Phishing
- Credentials Database Theft

Banking Trojans

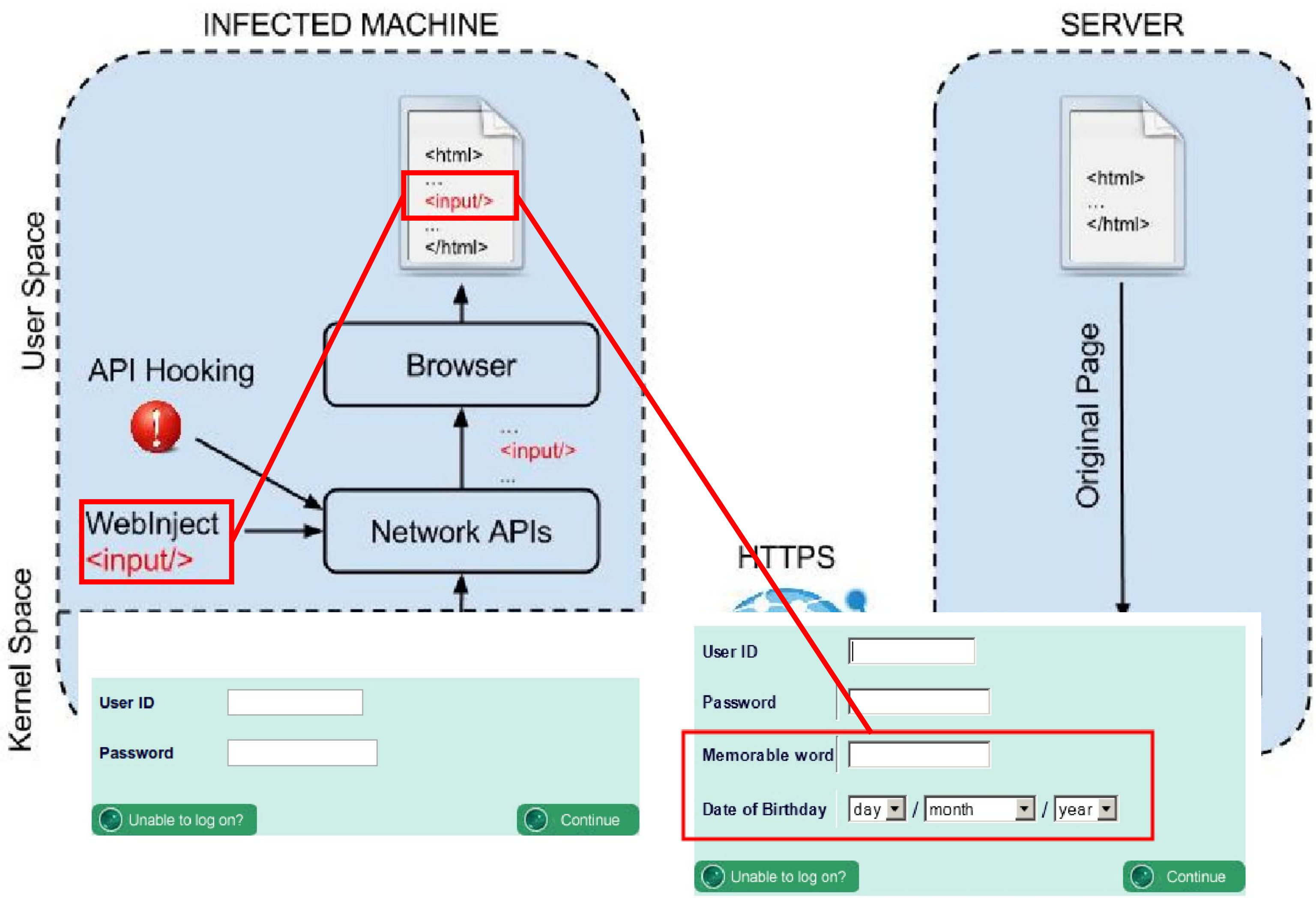
Malware that aims at stealing banking credentials in order to perform online financial frauds:

- ZeuS (2007+)
- SpyEye (2011+)
- Citadel (2012+)
- Carberp (2014+)

Anatomy of a Fraud

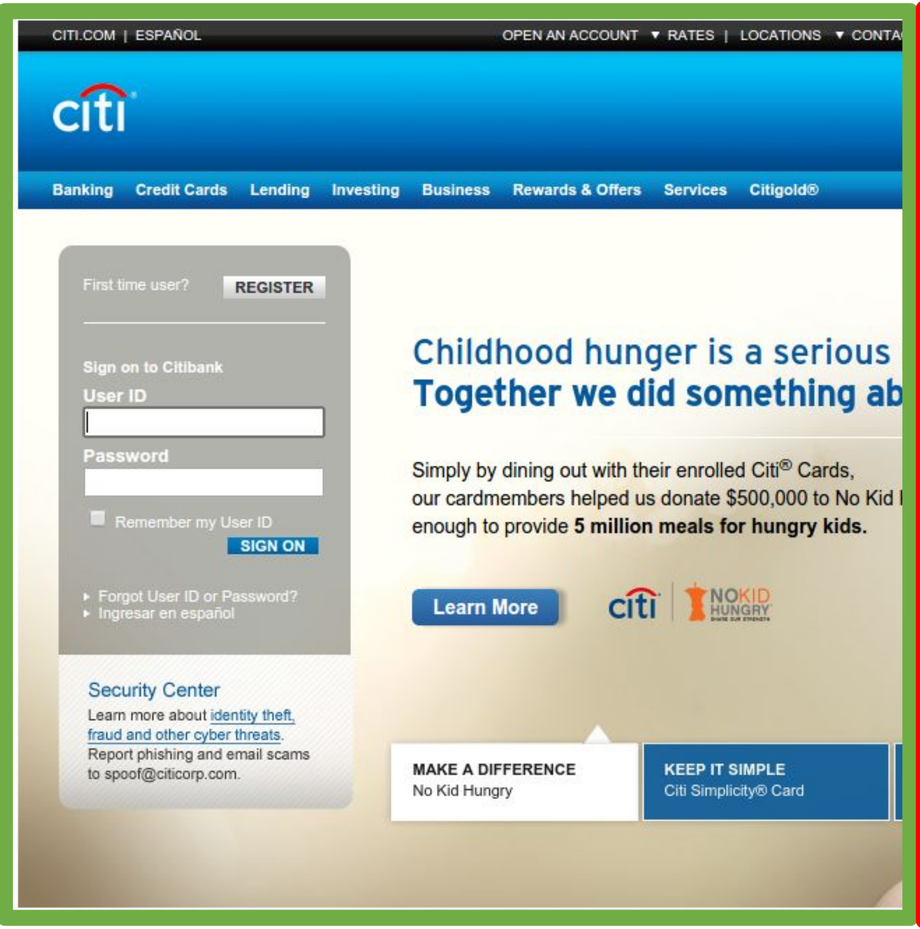


Web Injections



Effect of a web injection

Effect of a web injection



Banking Fraud Detection: Challenges

Internet banking **frauds** are **difficult to analyze and detect**:

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

Existing approaches & market offer

Fraud detection is a wide research topic

- Main focus: **credit cards**
  - Both in literature and in the market
- Most of the **existing approaches**:
  - **Black box:**
    - Instead, analysts need an explanation for the results
    - Tiring manual investigation and confirmation
  - **Not adaptive**:
    - CC frauds are assumed to be “always the same” across the world

BankSealer: 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 behavior** and its **evolution**

Publications

If you wish to check the publications during or after the talk:

Michele Carminati, Roberto Caron, Federico Maggi, Ilenia Epifani, Stefano Zanero, "BankSealer: An Online Banking Fraud Analysis and Decision Support System", in IFIP SEC 2014

Michele Carminati, Roberto Caron, Federico Maggi, Ilenia Epifani, Stefano Zanero, "BankSealer: A decision support system for online banking fraud analysis and investigation", Computers & Security, vol. 53, Sept. 2015, pp. 175–186

Original Dataset

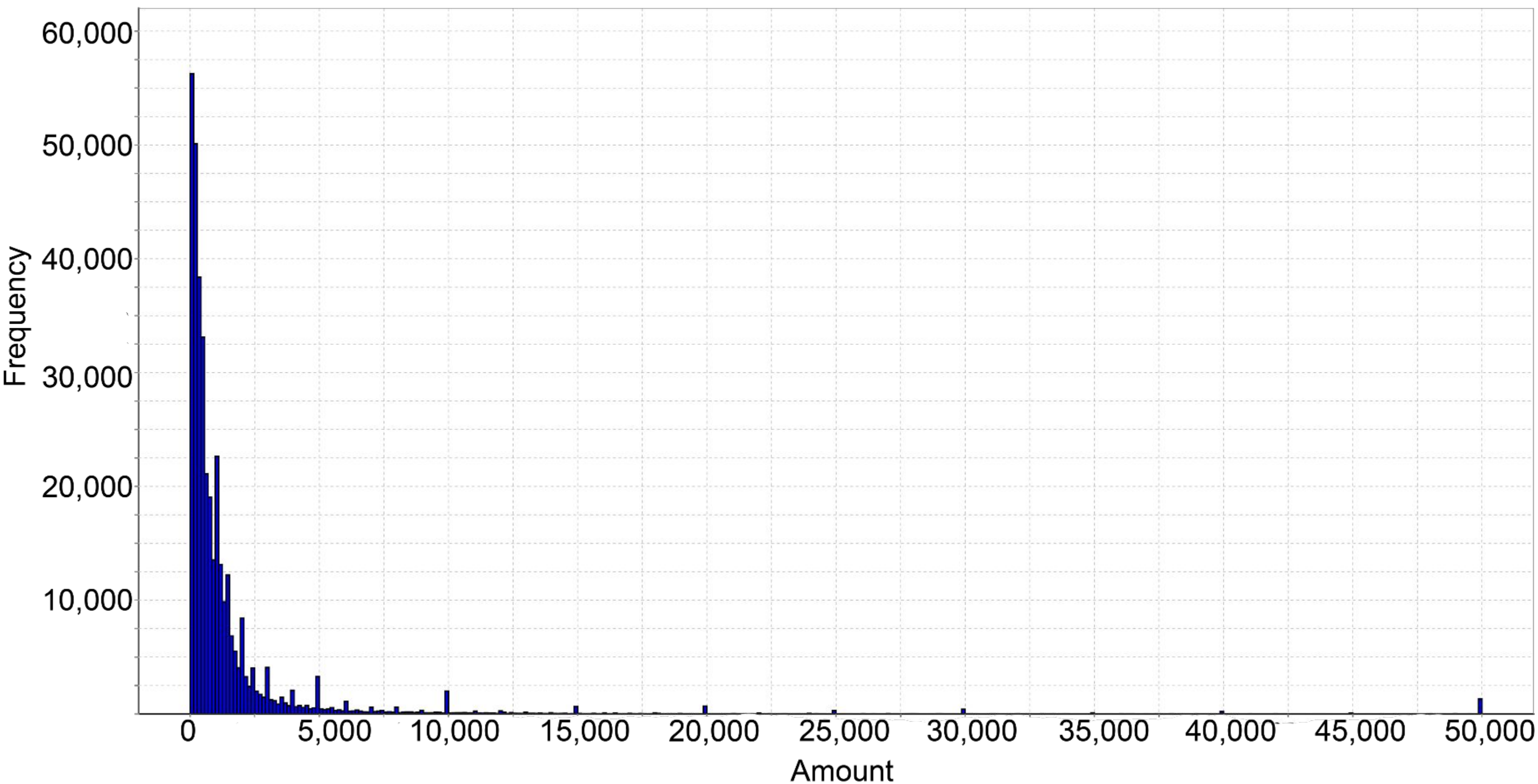
- Three months, one of the **largest banks in Italy**
- **Skewed** and **unbalanced** distribution of the attribute values
- High **cardinality**
- Majority of users perform only **few transactions**

	# Transactions	# Users
Bank Transfers	371,137	47,650
Prepaid phone	54,141	16,093
Debit cards	34,986	8,415

Anonymized Dataset

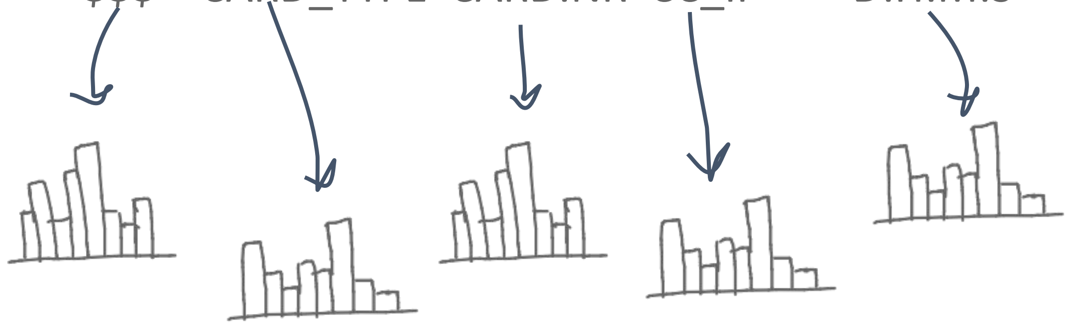
IP	Timestamp	Type	Amount	User ID	IBAN		Country
3d64e9f4a188aa034659d1409f90456a	08/feb/2013 21:06:30	Giroconto	20000	dcfc15d4d65e05ebafde6ac9383062aa	e6c2a617f55090de28a24c67dfbedf40	✔	IT
99ca402ce2299ecd72e2ebe269b5d35f	06/feb/2013 19:28:26	Bonifici per detrazione fiscale	9900	4a4ee6e2ac1b17e20958ad7a8221c1b4	11e6c83f92f065b07763f0beb451dd1a	✔	IT
ef7478c757b1fbe74bfbc21d1210051d	06/feb/2013 20:07:26	Giroconto	3000	e5d50d08b275cd98fb853ae6351c9fb9	11df01951ac12b1809feb39b81853fcf	✔	IT
91dfd27a9e1352963f8442edaa477c82	15/feb/2013 14:28:10	Bonifici per detrazione fiscale	22000	aa838ae340aa33fb774164f9a9985194	5ca5ba919c59ff70c5d522cc5fc97b61	✔	IT
86419f50fbda2742c1dba87cd3429470	28/feb/2013 17:46:46	Giroconto	12000	492ee7251c425c36c82cd3241c563f79	484fe271f1804b3e4291a537bb65279a	✔	IT
dd5d85da0532104875e18e4e32bc152c	08/feb/2013 16:11:24	Bonifici per detrazione fiscale	3863.29	a6f088c1fae1085def1308e532082cb7	73b5047423c9dad3b2601c929e2cece7	✔	IT
cd002daddde353900cc24e4ffc3b235c	21/feb/2013 18:40:02	Bonifici per detrazione fiscale	5643	a7b7a36b2769a1be86d1a544b67007a9	2626bfbc3376dababb639201c9b8ff67	✔	IT
0bdda3afaf28f049483d89d53f021c11	25/feb/2013 09:36:12	Bonifici Italia e SEPA	31000	be2b61118c081429cfbbc0c3d948743b	831687c224f781f106604f984e14f414	✔	IT
2aeddb8850ae946914285eb3bcd28d55	04/feb/2013 16:42:53	Giroconto	10000	41efb45d969e9511b7df6504840cc572	40c200429a2c2a4c7268b3300681e5e3	✔	IT
70c765c7265d92f96a05d91eebb4eb64	28/feb/2013 19:04:37	Bonifici Italia e SEPA	6529.6	8b7ed02e24a297a7ad7b91d28a5b35e1	3ada9624925ed42838bd4b8fab9eae81	✗	IT
d00d1939b4f71eaa199a57fff9cf0c19	20/feb/2013 14:59:10	Bonifici Italia e SEPA	50000	9bc3d0e6065284891a42ce6f9d828c38	65ecb9d1169b23049ec018d31c27af0a	✗	IT
2c2c6f325c547ee1fb0efc01475bc7d6	14/feb/2013 09:17:53	Bonifici Italia e SEPA	50000	f2a7341750c1cc6dc8bea45185a7fe26	60414014d030aa24b4cef90c32fac61f	✔	PT
ba3664bb7ebbf9e8bf4ac0664d65e239	10/feb/2013 19:21:11	Bonifici Italia e SEPA	20000	2a17ed71d9e2c82f39e174e424bf7eb9	e9987193889c72a6dcb94bbd47e35699	✗	IT
d7ab9d7839eb60ff6e606c496e1c848a	08/feb/2013 19:23:46	Bonifici Italia e SEPA	25266.8	435b8226966d2fb40d52bafd6aaa8a93	92a91621f34b668401e8c26050f4e0c6	✗	IT
6da1465327246224216c1c929c339c6f	18/feb/2013 15:09:11	Bonifici Italia e SEPA	50000	f2a7341750c1cc6dc8bea45185a7fe26	60414014d030aa24b4cef90c32fac61f	✔	PT

Skewed data example

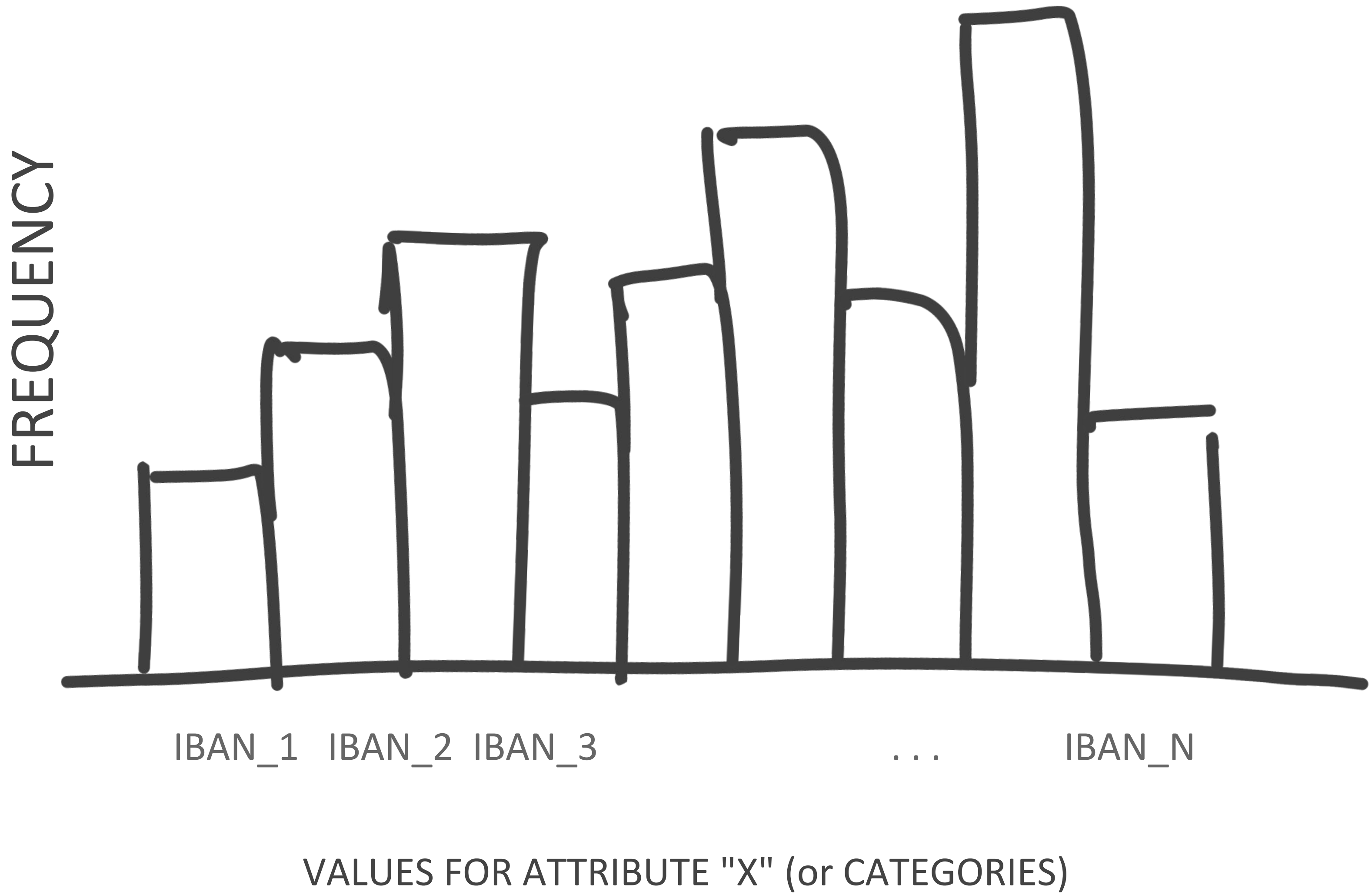


Attributes

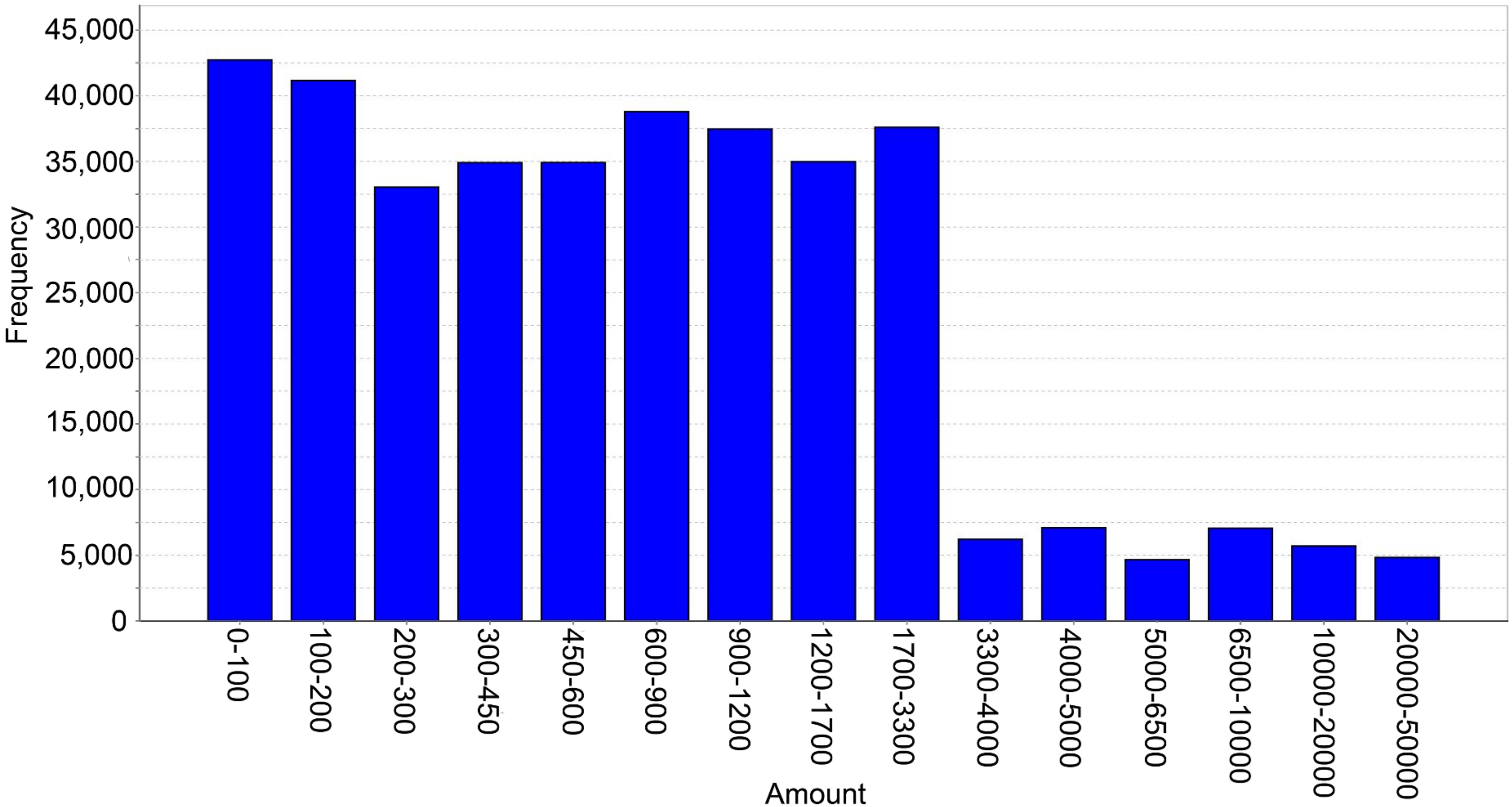
BANK TRANSFERS	$$$	CC.IP	IP	IBAN	CC_IBAN	D:H:M:S
PHONE RECAHRGES	$$$	CC.IP	IP	OP.TEL	NUM.TEL	D:H:M:S
PREPAID CARDS	$$$	CARD_TYPE	CARD.NR	CC_IP	D:H:M:S	



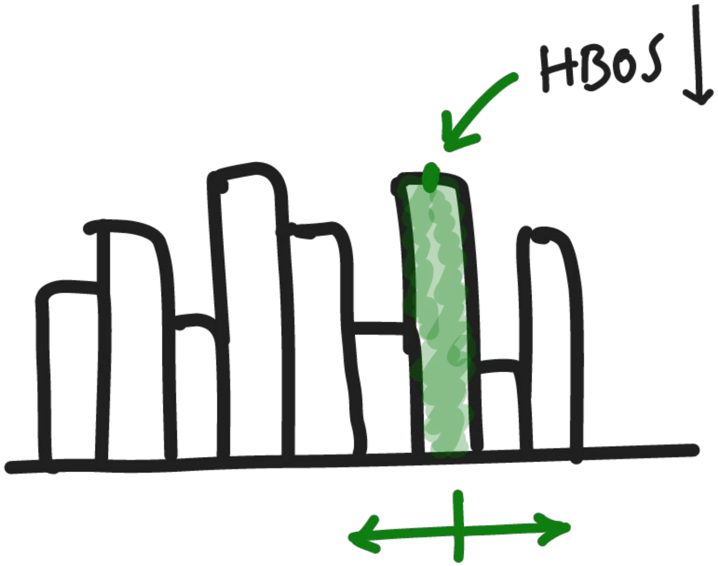
Model for Each Attribute

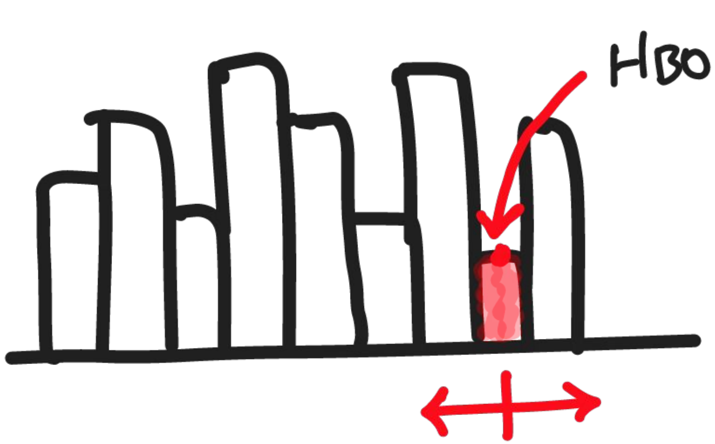


Model Example



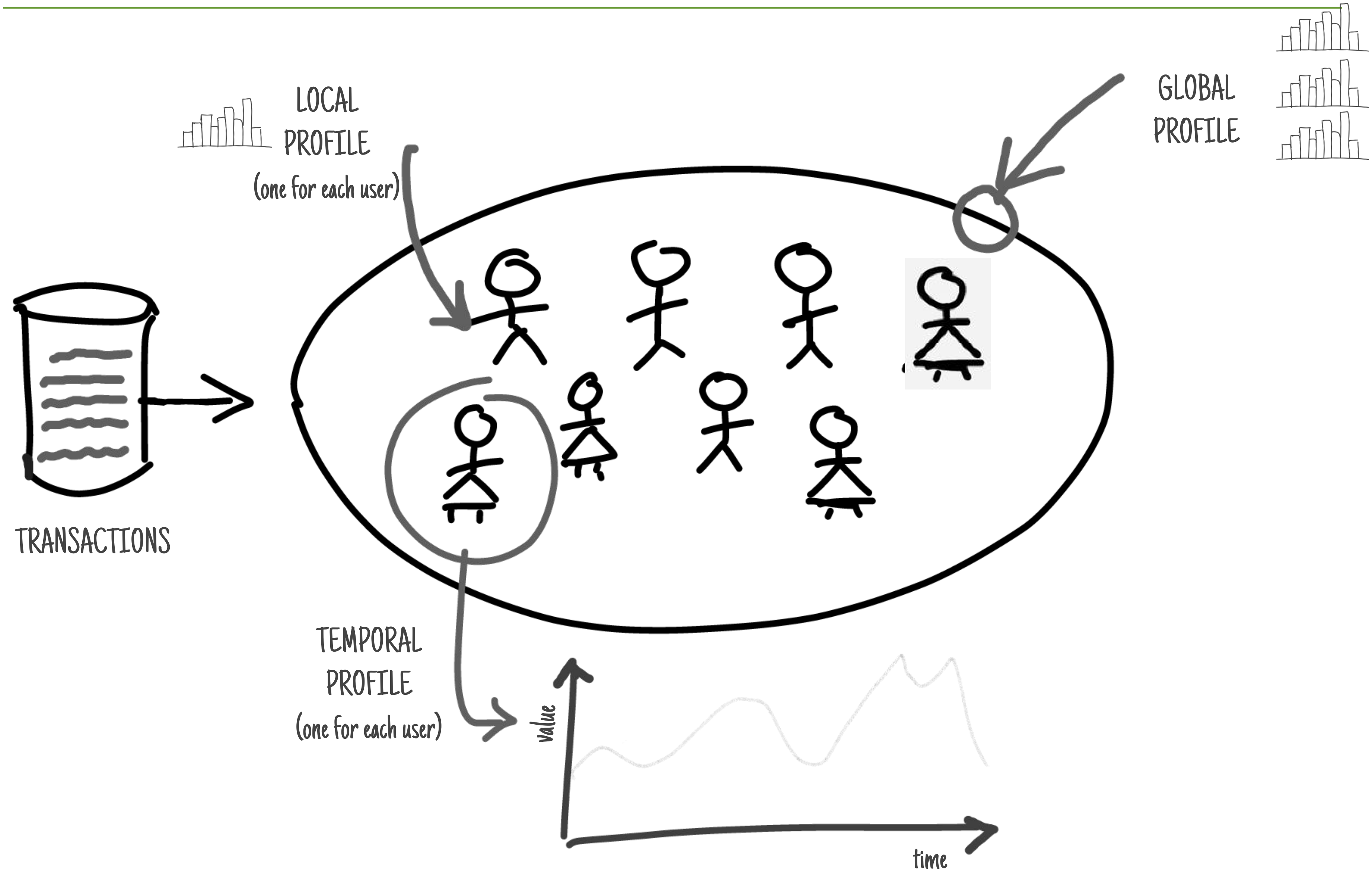
HBOS = Histogram Based Outlier Score



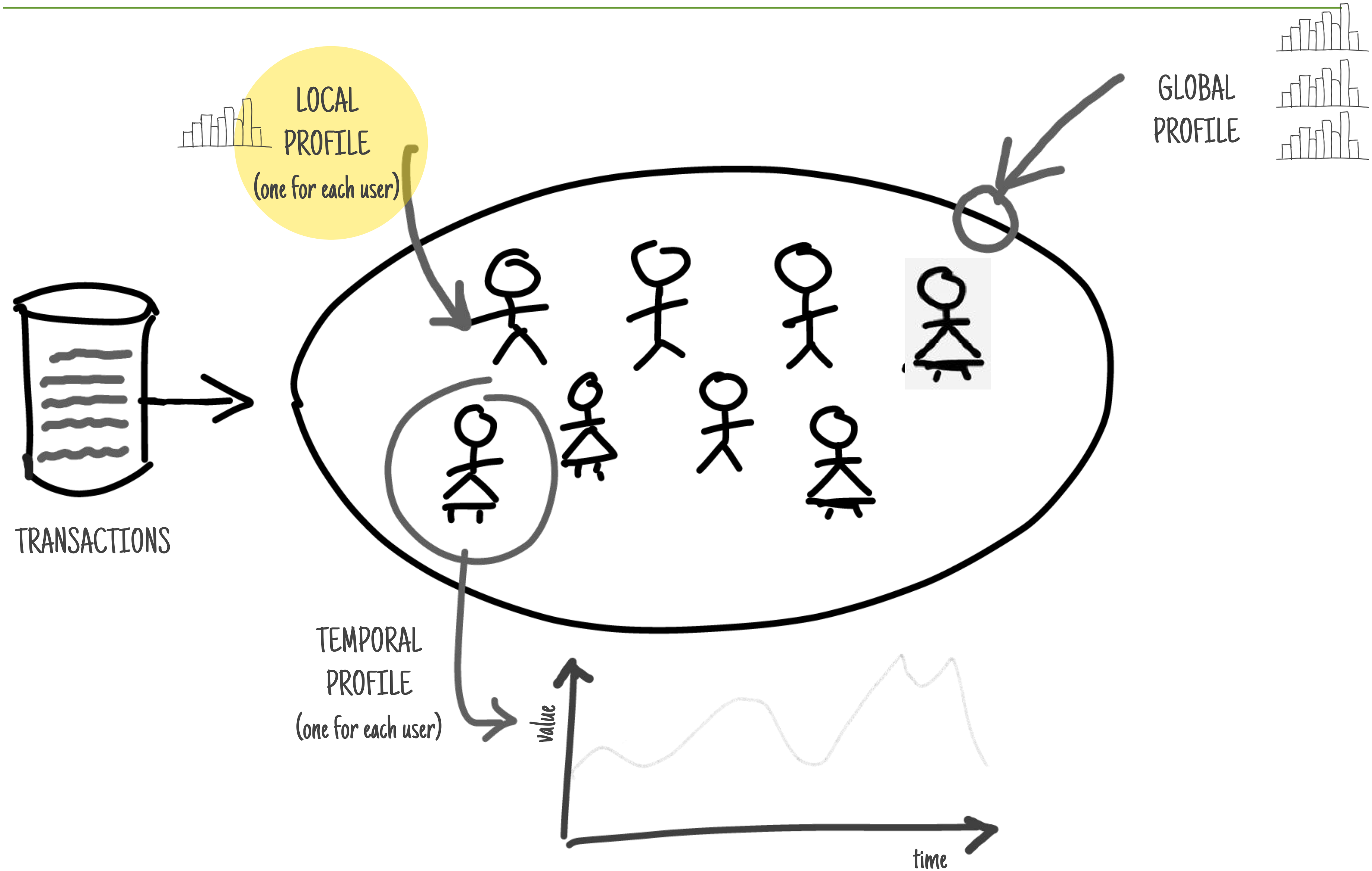


$$HBOS(t) = \sum_{0<i\leq d} w_i * \log \frac{1}{f(t_i)}; \qquad \sum_{0<i\leq d} w_i = 1$$

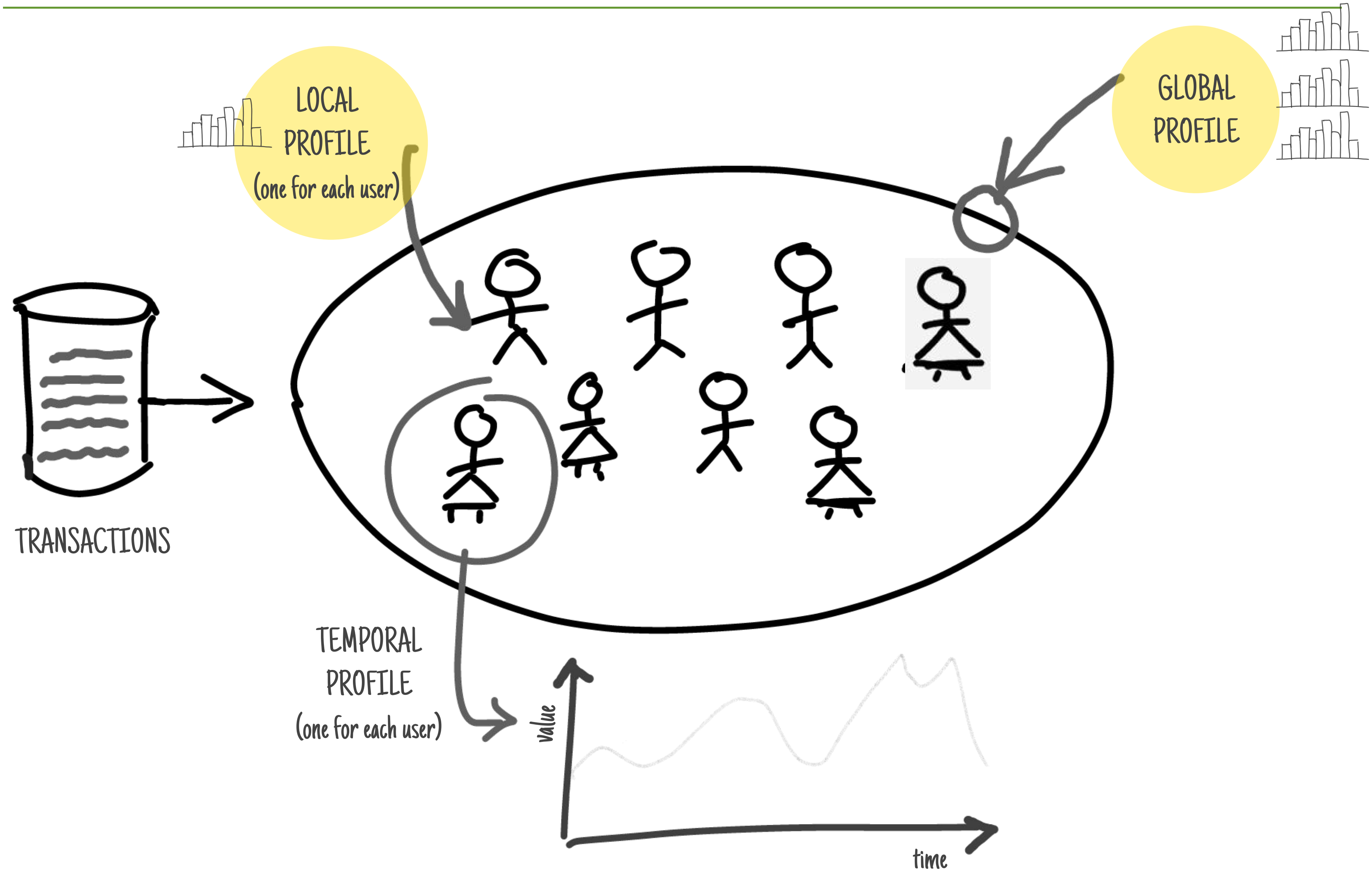
Putting it all together



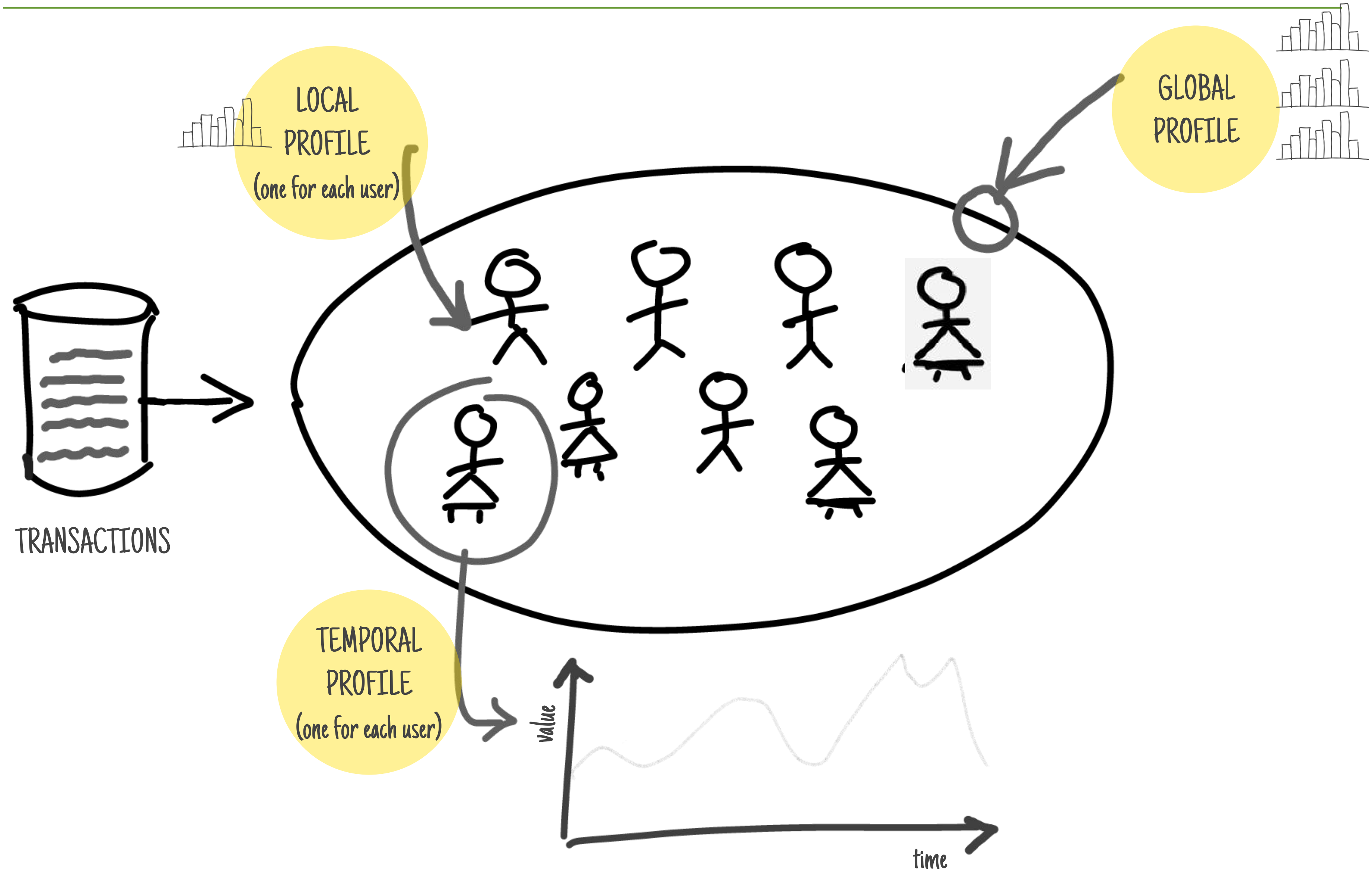
Putting it all together



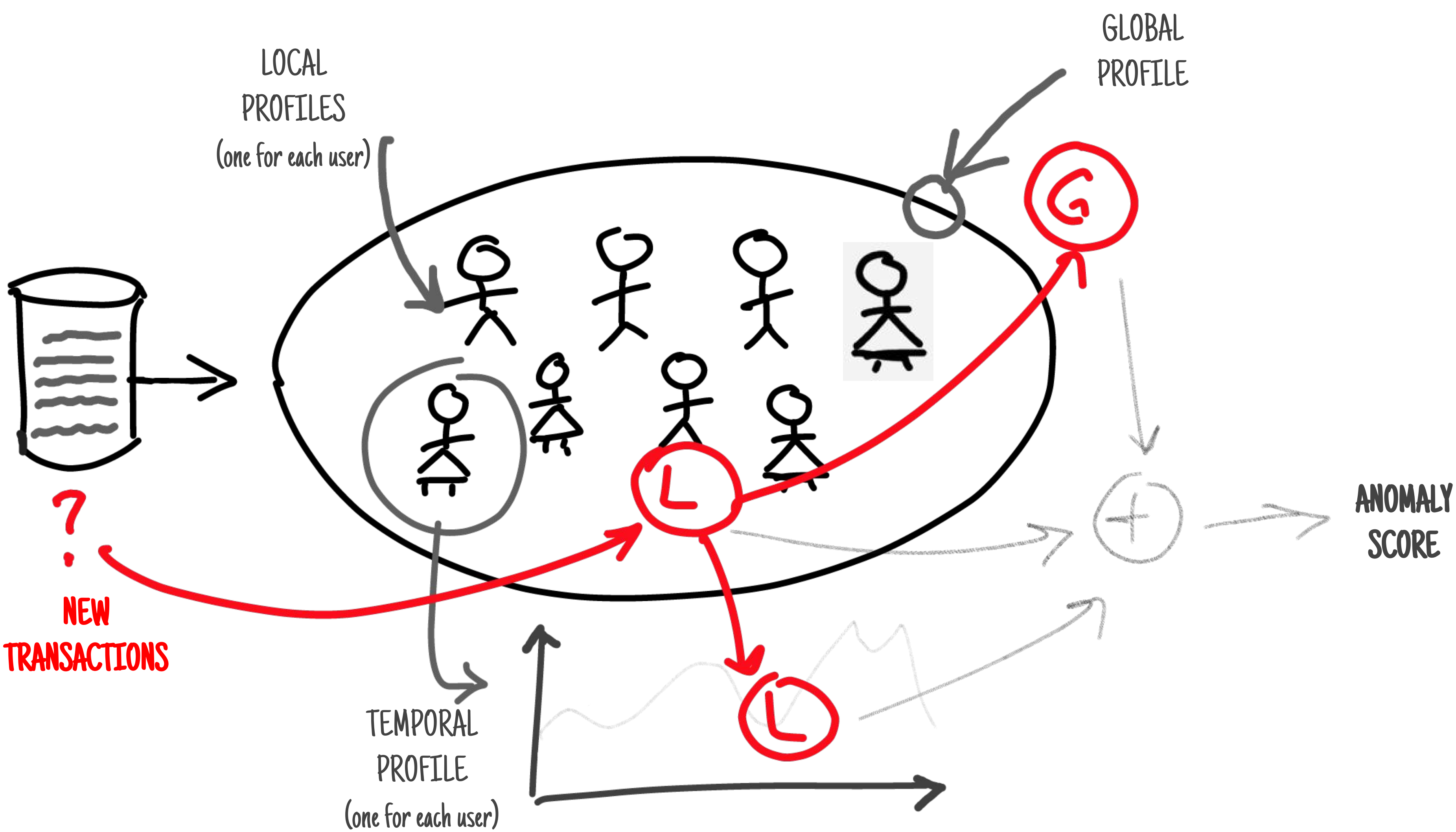
Putting it all together



Putting it all together

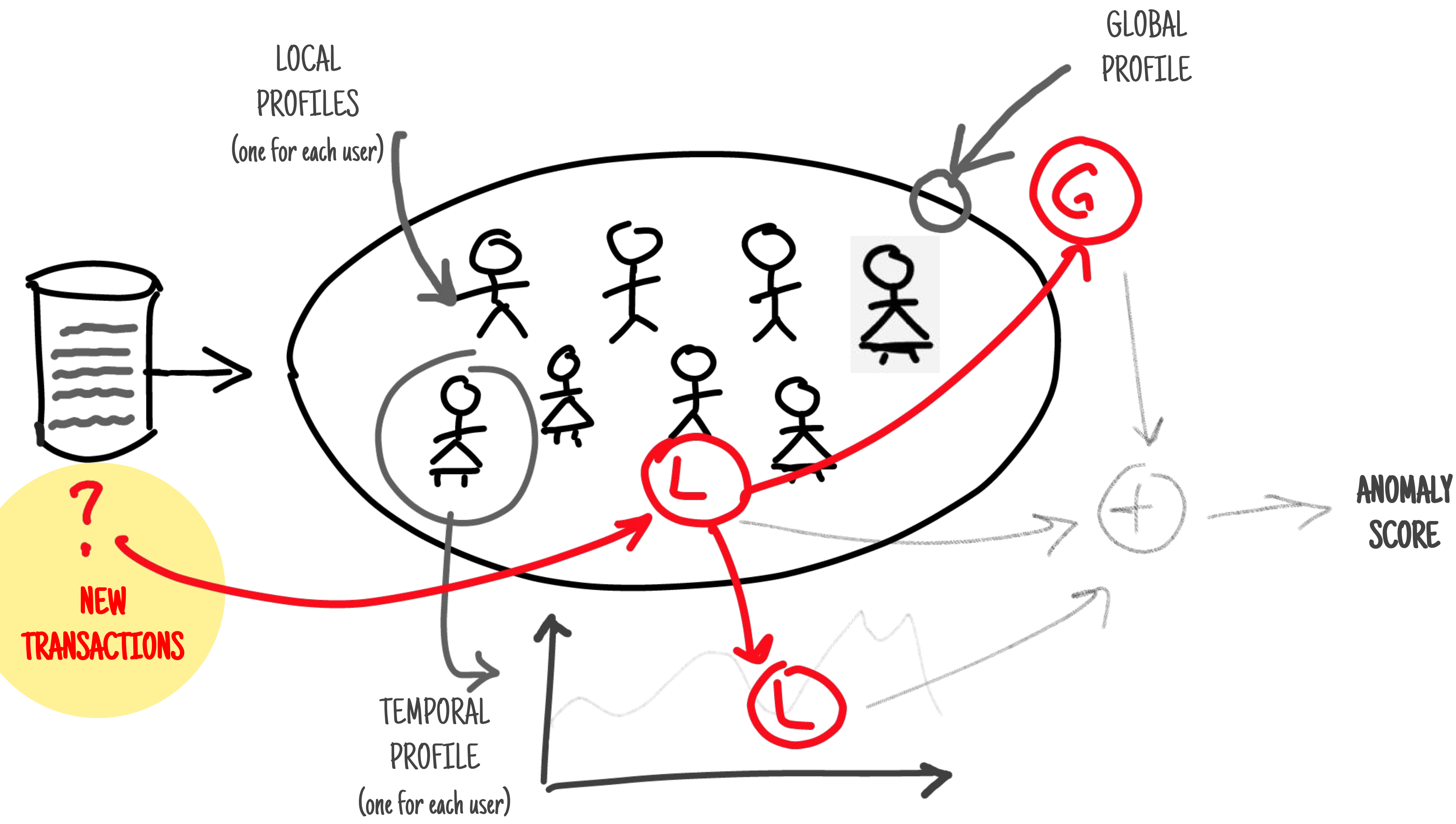


Anomalies



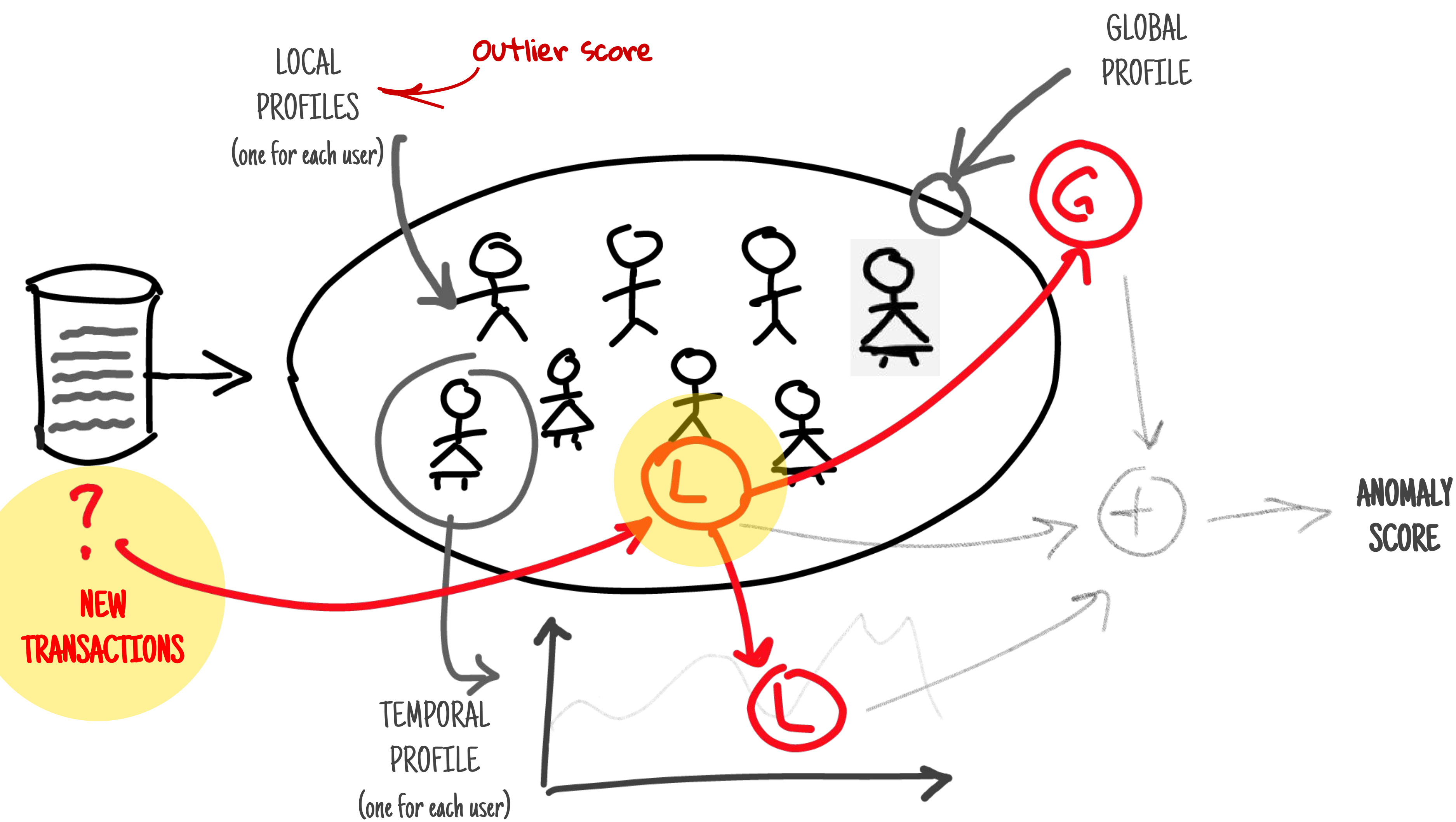
LOCAL
PROFILES
(one for each user)
GLOBAL
PROFILE
G
?
NEW
TRANSACTIONS
L
L
ANOMALY
SCORE
TEMPORAL
PROFILE
(one for each user)

Anomalies



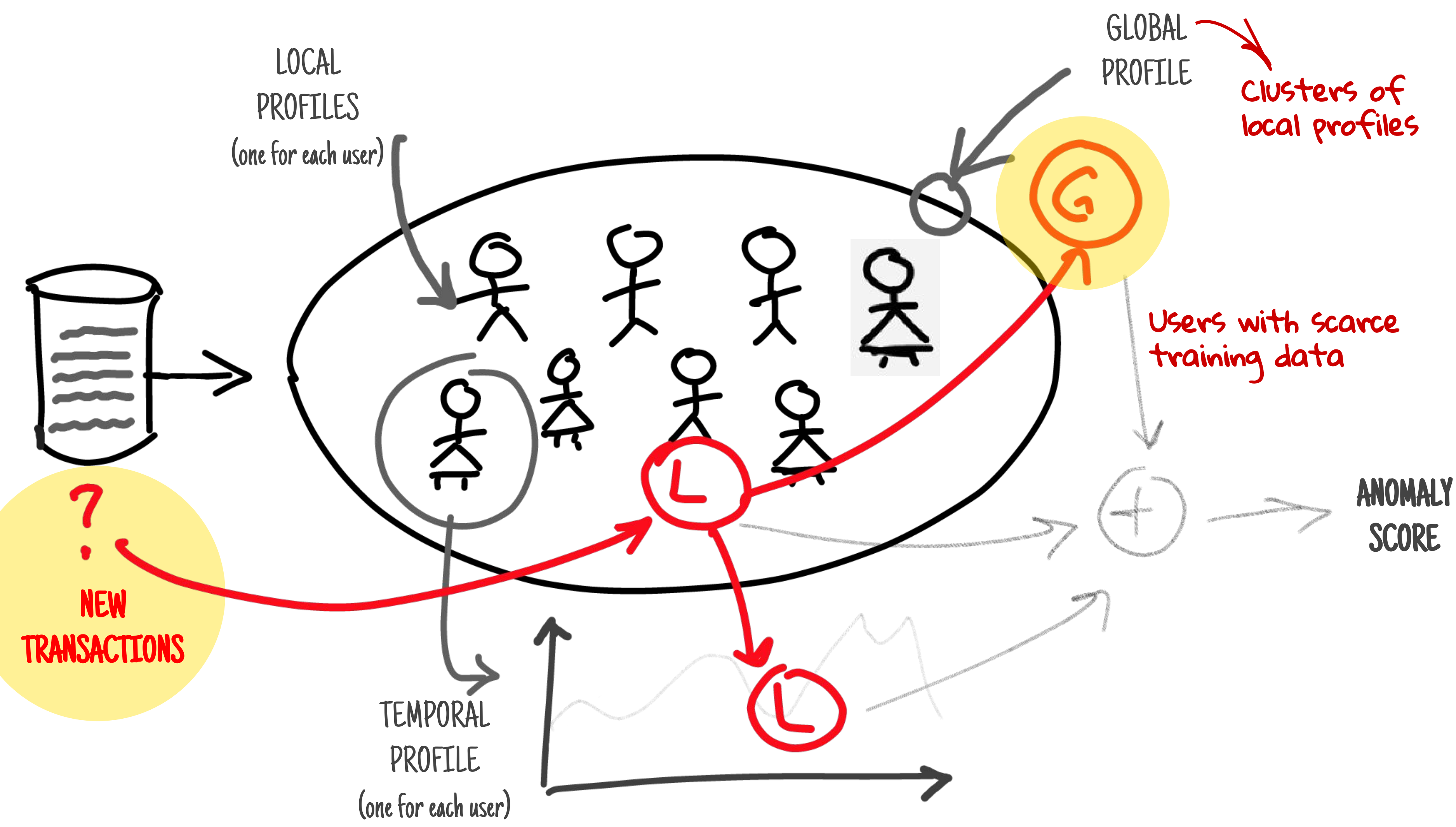
LOCAL
PROFILES
(one for each user)
GLOBAL
PROFILE
G
L
L
?
NEW
TRANSACTIONS
ANOMALY
SCORE
TEMPORAL
PROFILE
(one for each user)

Anomalies

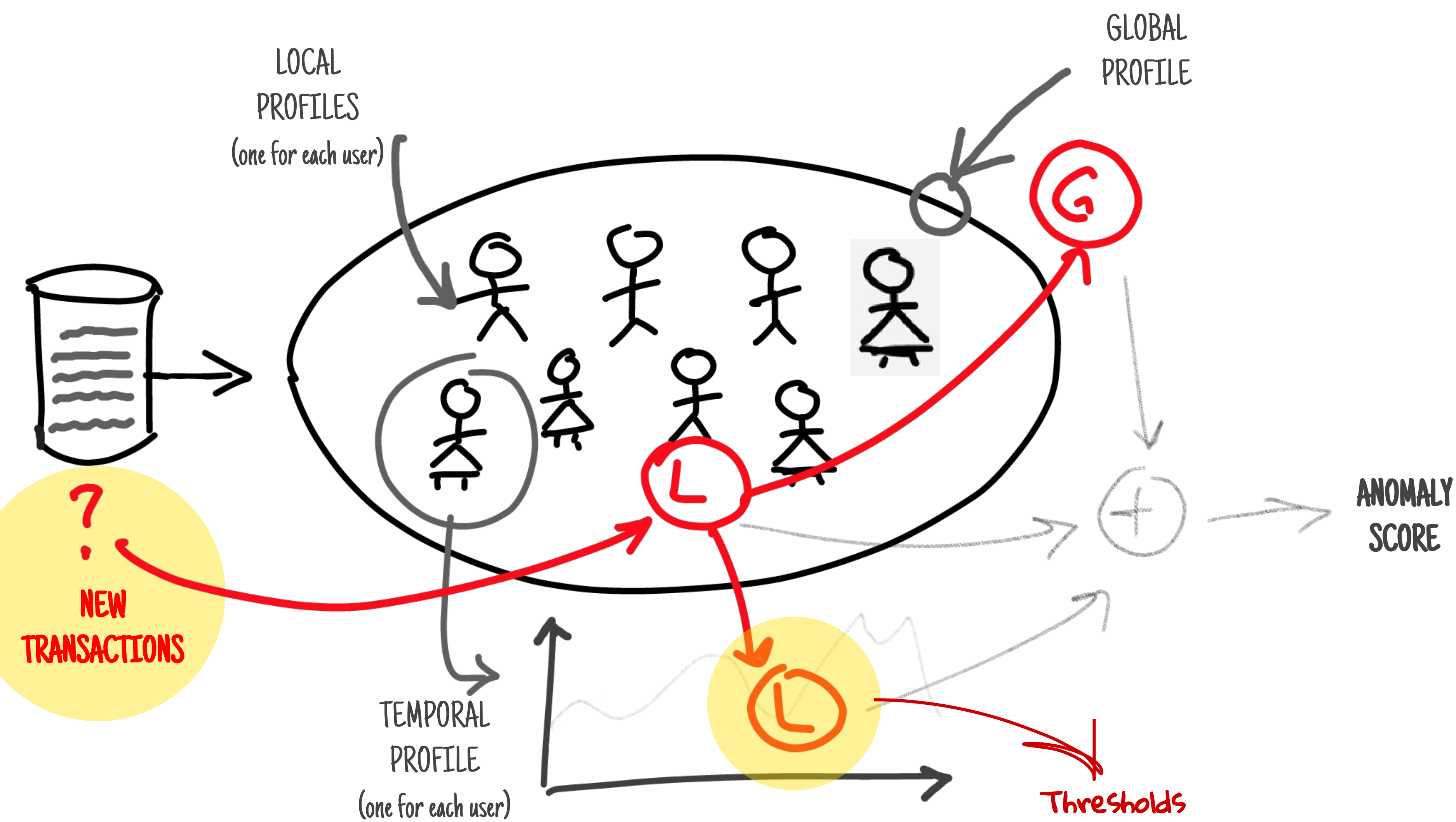


LOCAL
PROFILES
(one for each user)
Outlier score
GLOBAL
PROFILE
G
L
?
NEW
TRANSACTIONS
L
ANOMALY
SCORE
TEMPORAL
PROFILE
(one for each user)

Anomalies

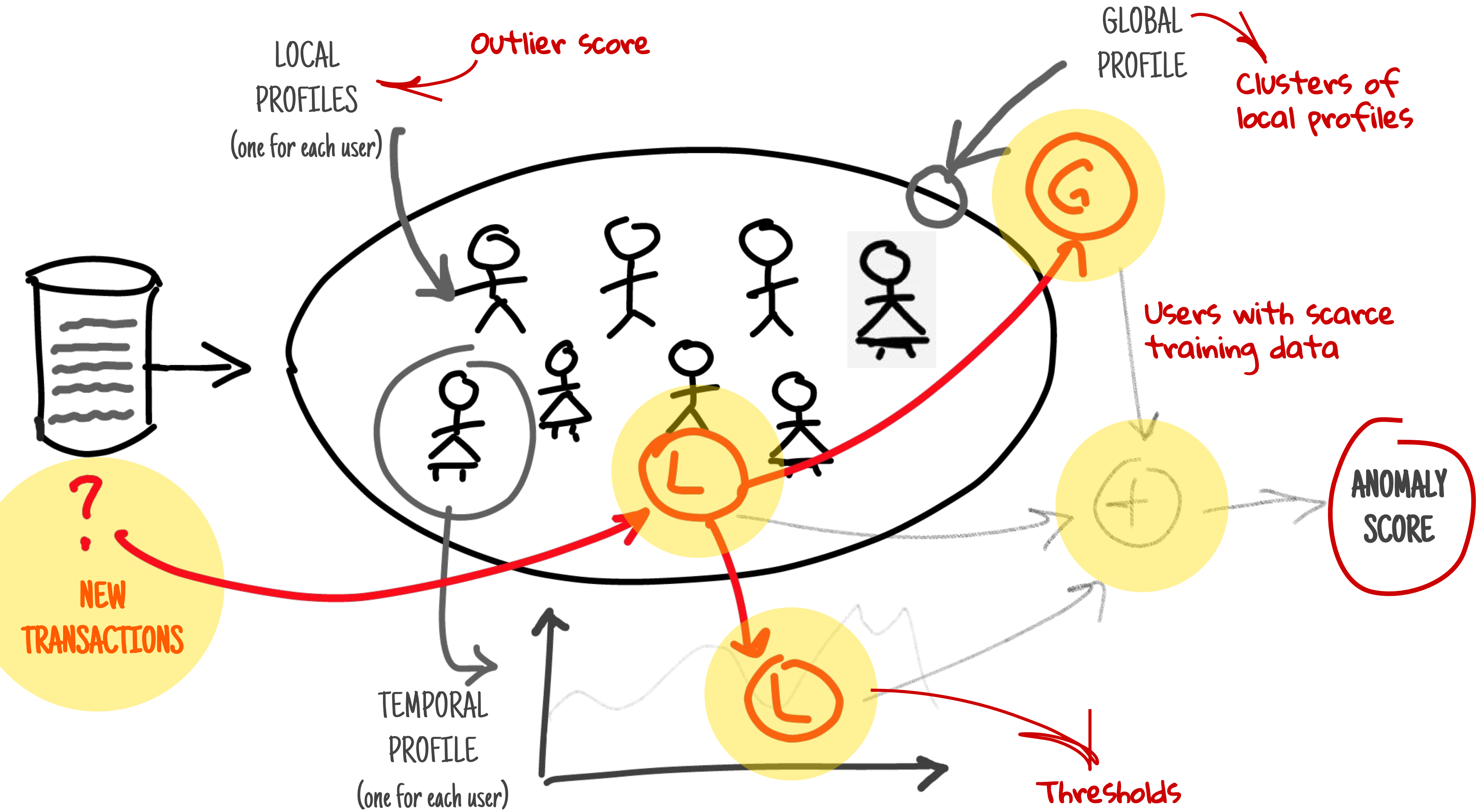

LOCAL
PROFILES
(one for each user)
GLOBAL
PROFILE
Clusters of
local profiles
G
Users with scarce
training data
L
?
NEW
TRANSACTIONS
L
ANOMALY
SCORE
TEMPORAL
PROFILE
(one for each user)

Anomalies



LOCAL PROFILES (one for each user)
GLOBAL PROFILE
G
L
? NEW TRANSACTIONS
ANOMALY SCORE
TEMPORAL PROFILE (one for each user)
L
Thresholds

Anomalies



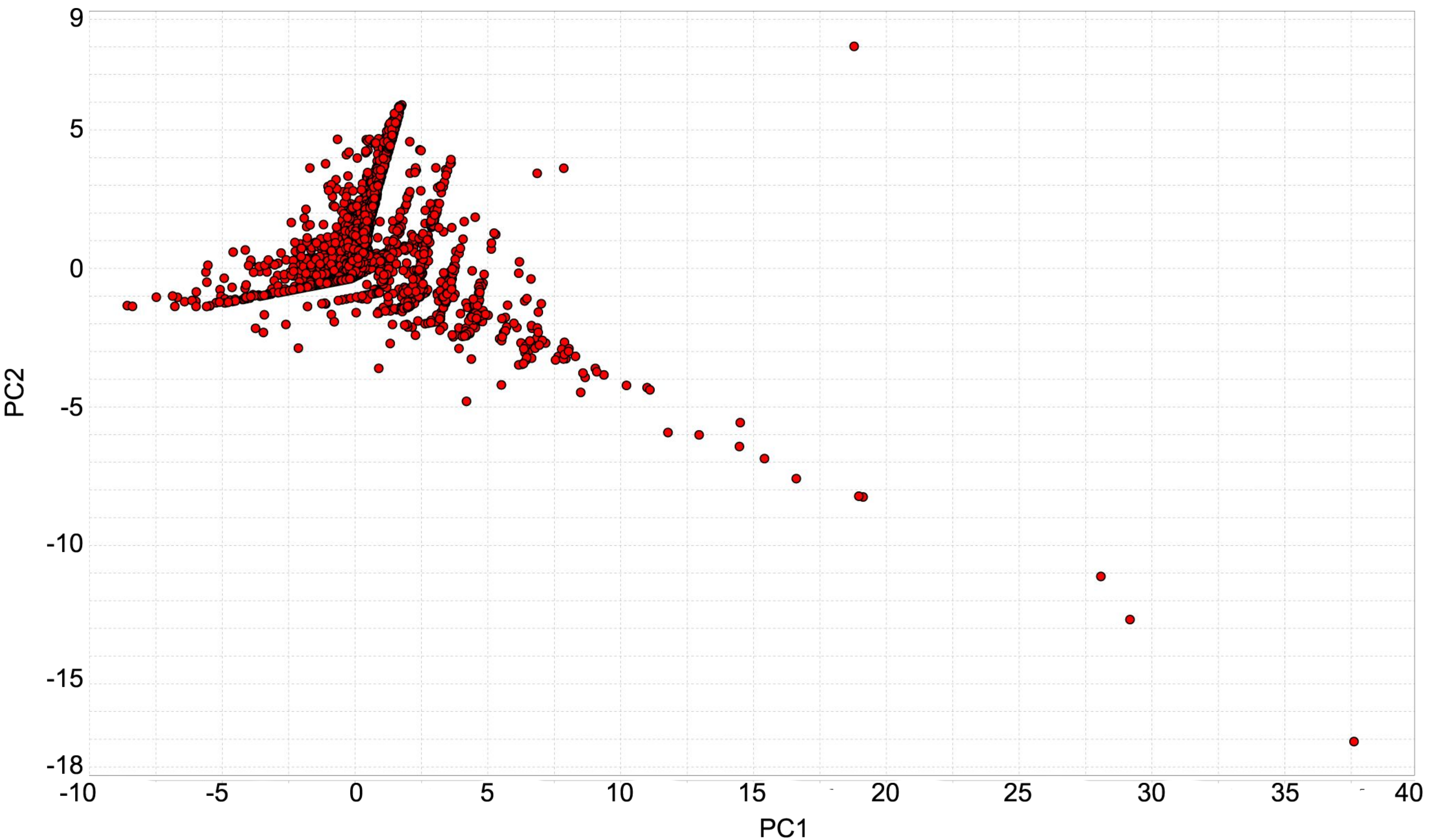
LOCAL
PROFILES
(one for each user)
Outlier score
GLOBAL
PROFILE
Clusters of
local profiles
Users with scarce
training data
ANOMALY
SCORE
?
NEW
TRANSACTIONS
TEMPORAL
PROFILE
(one for each user)
Thresholds

Users with scarce data: Global Profile

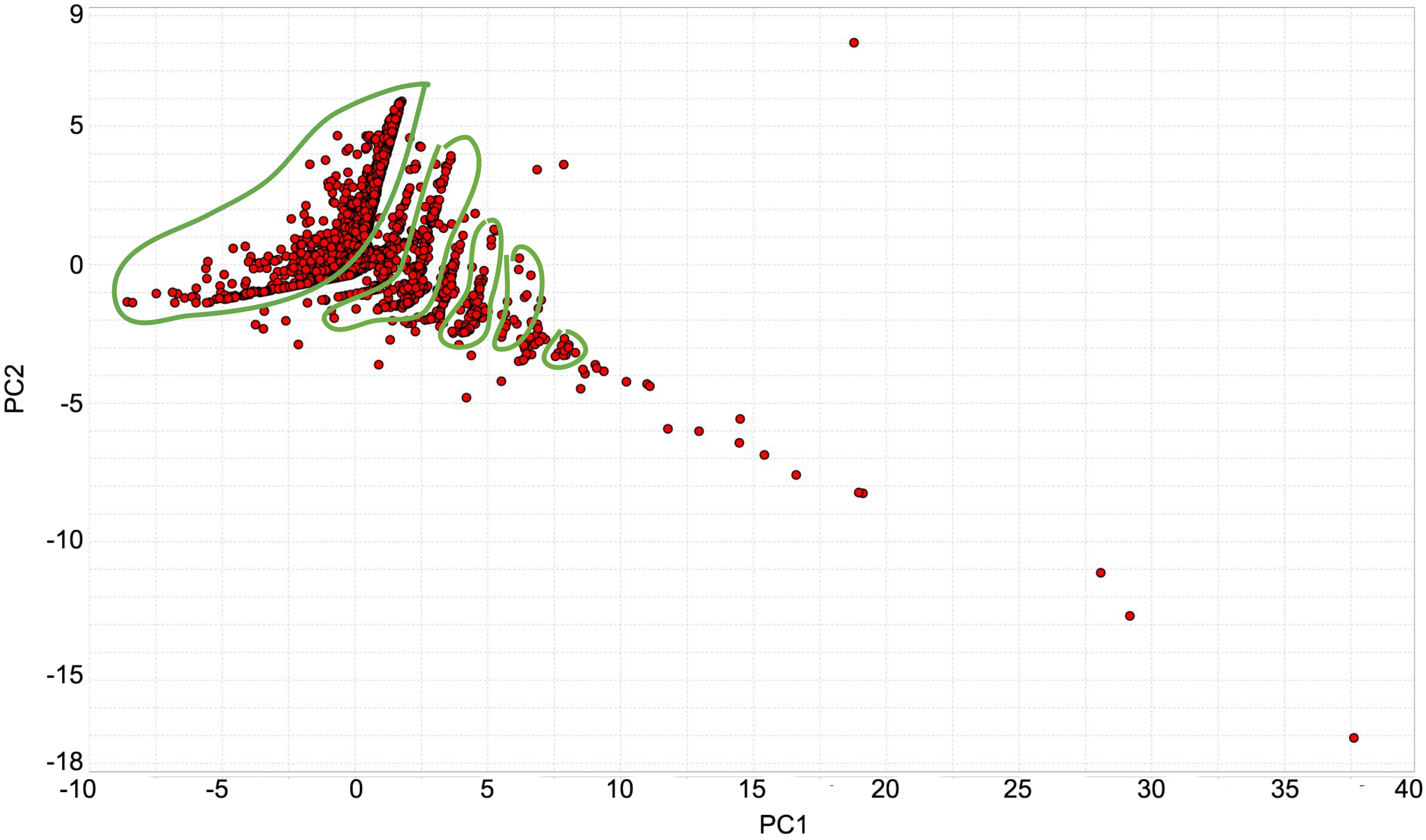
Two phases:

1. **Clustering:** find groups of *similar* users
   a. **Algorithm:** incremental DBSCAN
   b. **Distance function:** Mahalanobis
2. **Anomaly score:** *distance* of a user from the large clusters

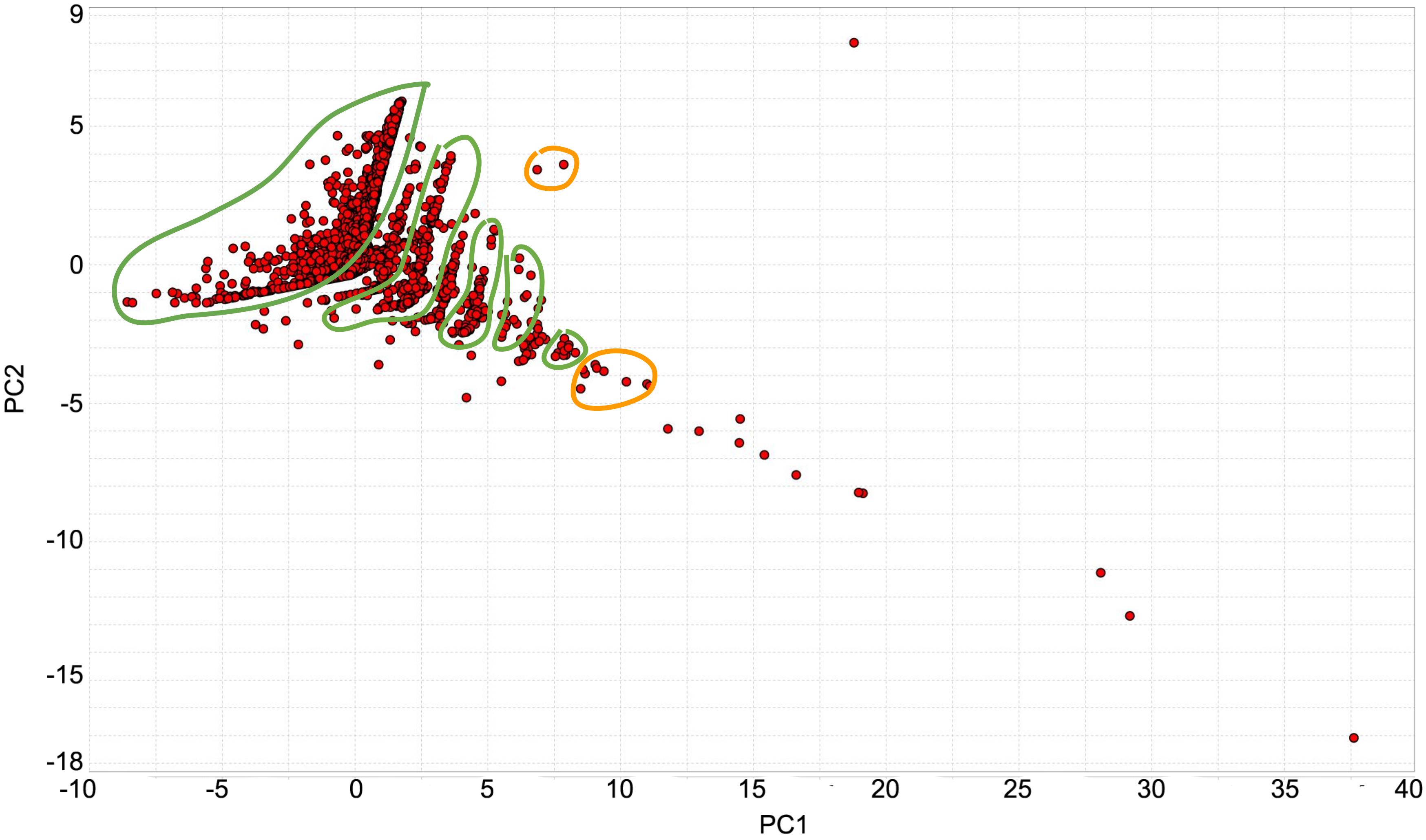
Majority of Users Behave Similarly



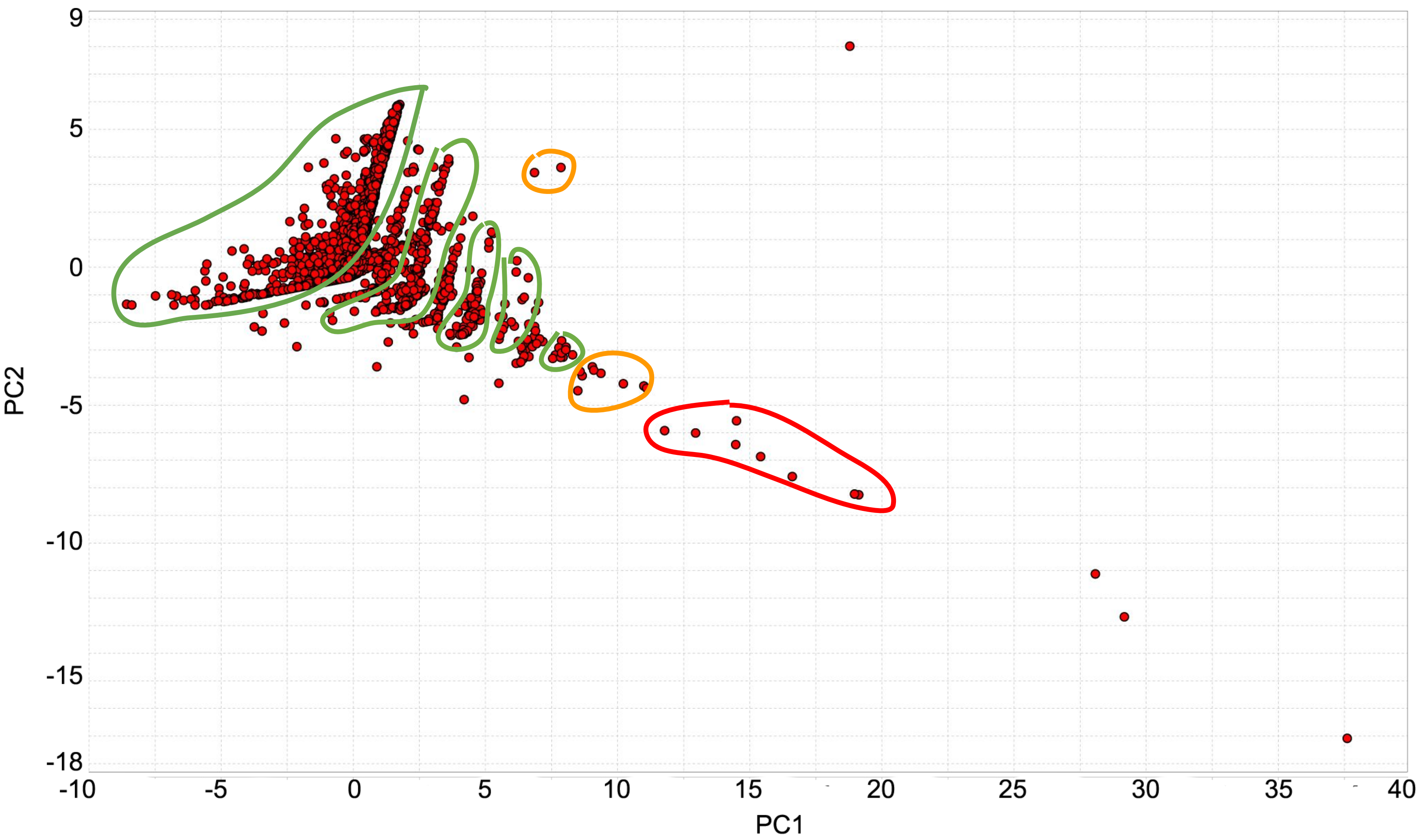
Majority of Users Behave Similarly



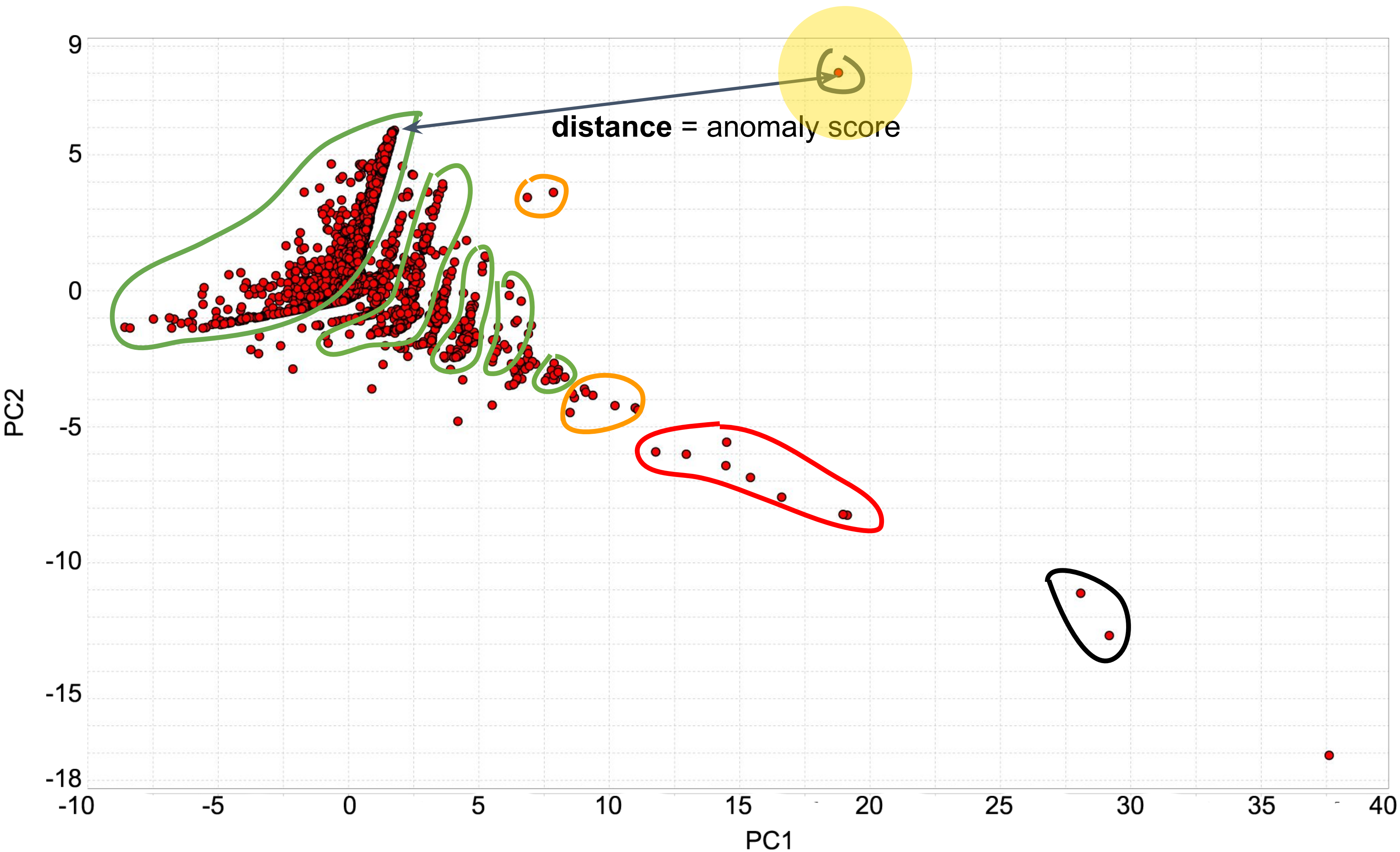
Majority of Users Behave Similarly



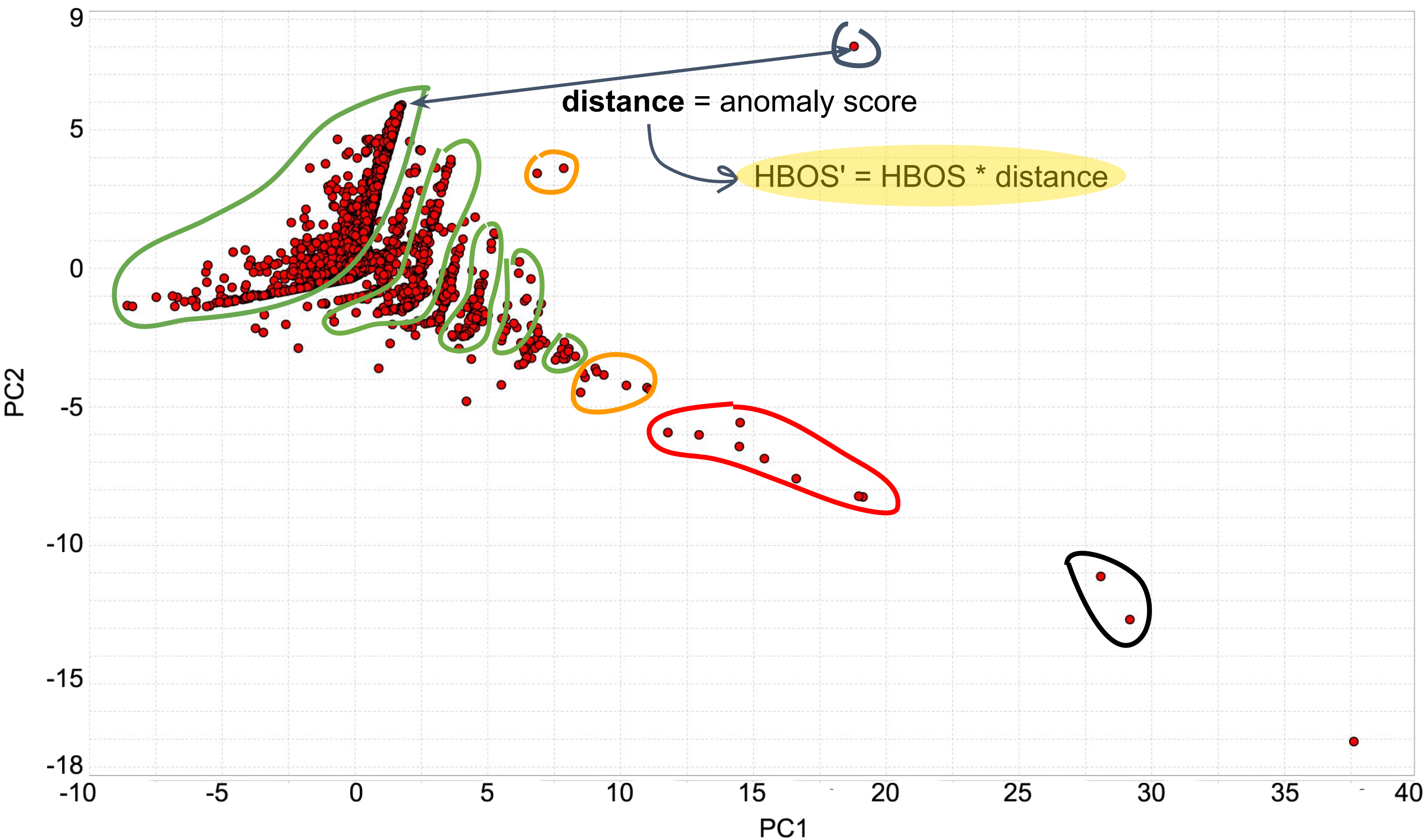
Majority of Users Behave Similarly



Majority of Users Behave Similarly

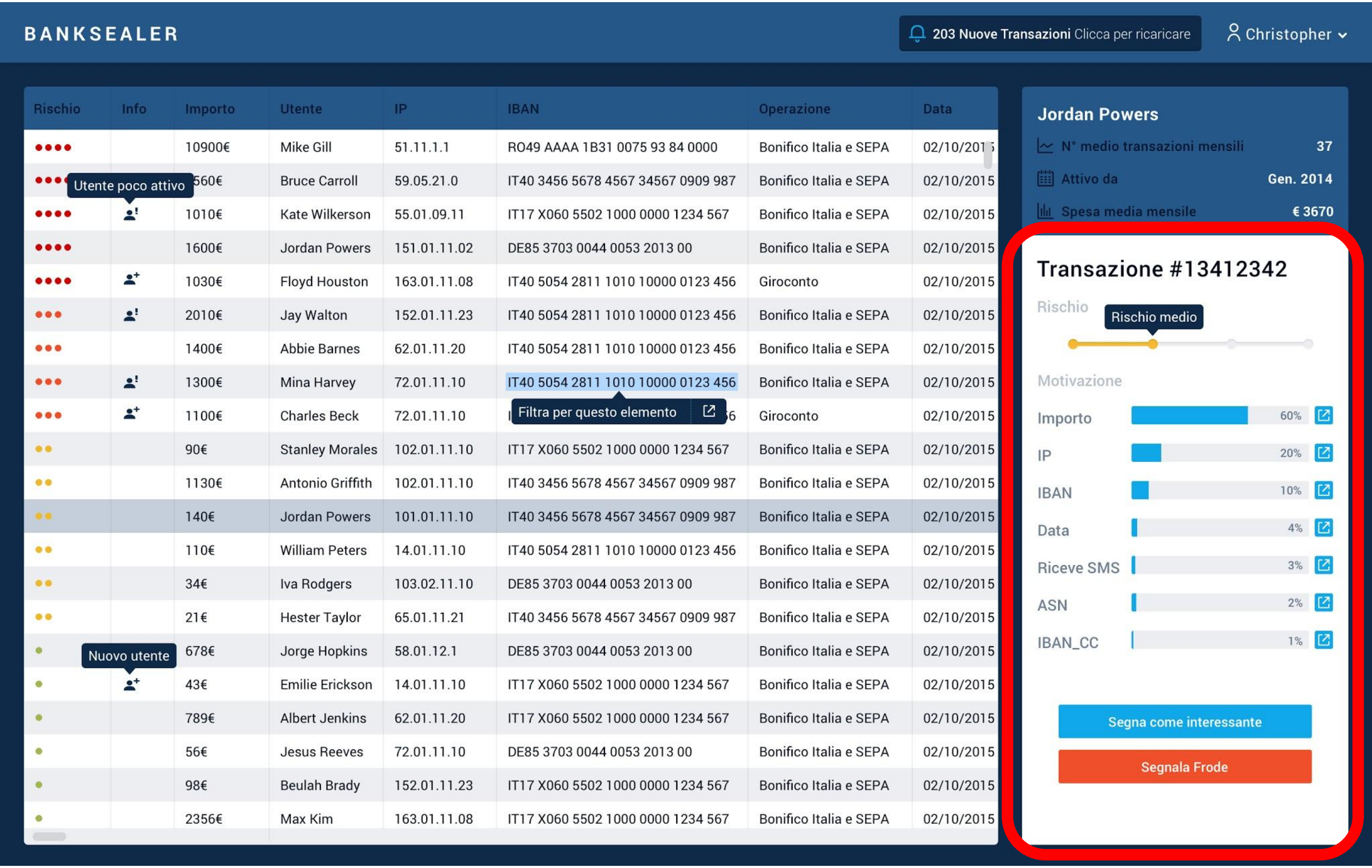


Majority of Users Behave Similarly

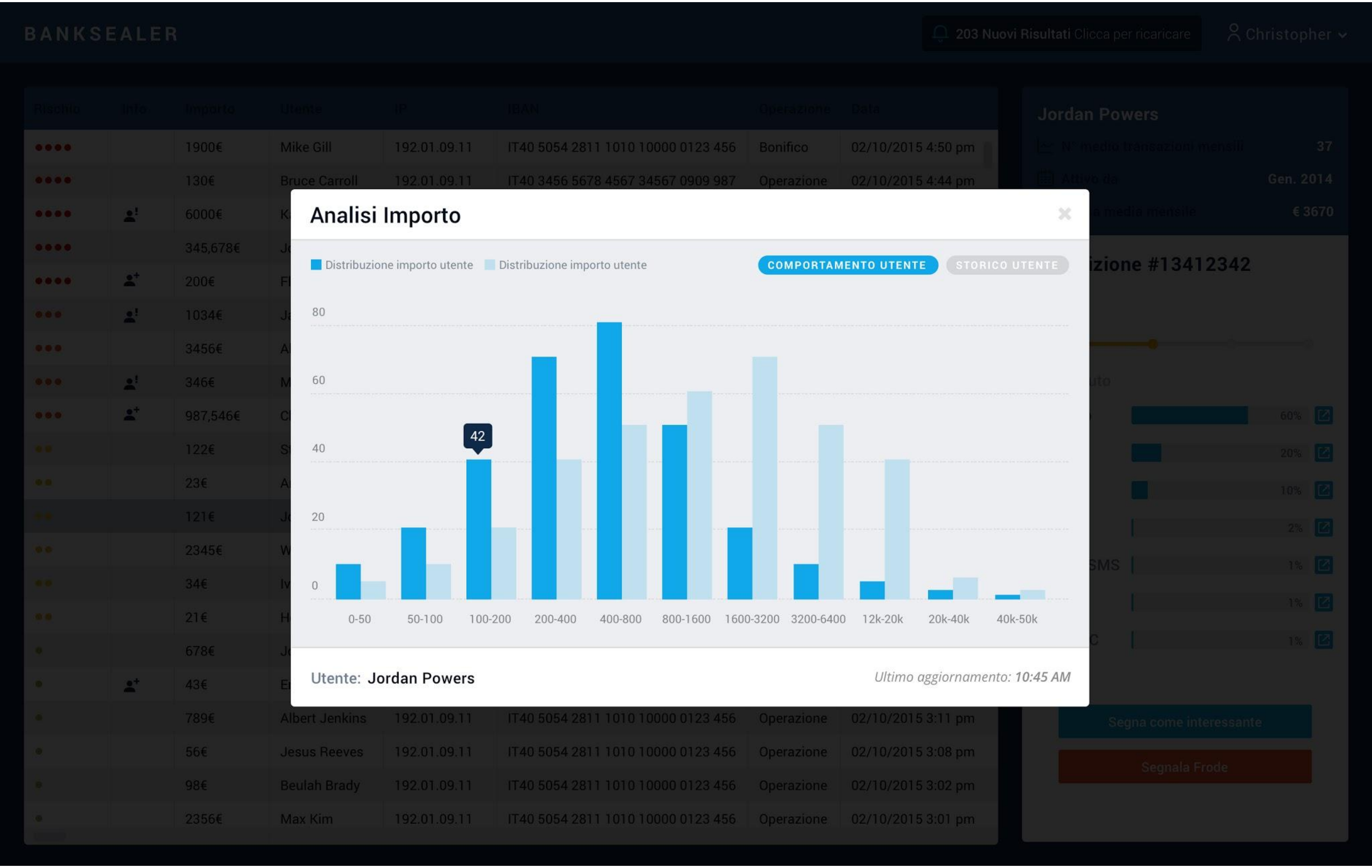


	index	IP	Timestamp	TipoOperazione	Importo	UserID	IBAN	Numero Conferma SMS	IBAN_CC	CC_ASN	HBOS locale	Undertrained	New User
1	92580	3d64e9f4a188aa034659d1409f90456a	08/feb/2013 21:06:30	Giroconto	20000	dcfc15d4d65e05ebafde6ac9383062aa	e6c2a617f55090de28a24c67dfbedf40	✓	IT	IT,8612	29.402365073	✓	✗
2	91133	99ca402ce2299ecd72e2ebe269b5d35f	06/feb/2013 [illegible]	Bonifici per detrazione [illegible]	9900	4a4ee6e2ac1b17e20958ad7a8221c1b4	[illegible]	✓	IT	IT,3269	27.0032775111	✓	✗
3	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	✗	✗
4	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	✗	✗
5	101355	86419f50fbda2742c1dba87cd3429470	28/feb/2013 17:40:46	Giroconto	12000	492ee7251c425c36c82cd3241c563f79	484fe271f1804b3e4291a537bb65279a	✓	IT	IT,44957	25.0995700682	✓	✗
6	92502	dd5d85da0532104875e18e4e32bc152c	08/feb/2013 16:11:24	Bonifici per detrazione fiscale	3863.29	a6f088c1fae1085def1308e532082cb7	73b5047423c9dad3b2601c929e2cece7	✓	IT	IT,30722	24.5660880611	✓	✗
7	99074	cd002daddde353900cc24e4ffc3b235c	21/feb/2013 18:40:02	Bonifici per detrazione fiscale	5643	a7b7a36b2769a1be86d1a544b67007a9	2626bfbc3376dababb639201c9b8ff67	✓	IT	IT,21056	24.4493640116	✓	✗
8	99827	0bdda3afaf28f049483d89d53f021c11	25/feb/2013 09:36:12	Bonifici Italia e SEPA	31000	be2b61118c081429cfbbc0c3d948743b	831687c224f781f106604f984e14f414	✓	IT	IT,12874	24.4175445119	✓	✗
9	89586	2aeddb8850ae946914285eb3bcd28d55	04/feb/2013 16:42:53	Giroconto	10000	41efb45d969e9511b7df6504840cc572	40c200429a2c2a4c7268b3300681e5e3	✓	IT	IT,12874	23.6879134642	✓	✗
10	101627	70c765c7265d92f96a05d91eebb4eb64	28/feb/2013 19:04:37	Bonifici Italia e SEPA	6529.6	8b7ed02e24a297a7ad7b91d28a5b35e1	3ada9624925ed42838bd4b8fab9eae81	✗	IT	IT,50809	23.6370584204	✓	✗
11	98401	d00d1939b4f71eaa199a57fff9cf0c19	20/feb/2013 14:59:10	Bonifici Italia e SEPA	50000	9bc3d0e6065284891a42ce6f9d828c38	65ecb9d1169b23049ec018d31c27af0a	✗	IT	IT,3269	23.5882551789	✓	✗
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	✗	IT	IT,3269	23.5233646465	✓	✗
14	92551	d7ab9d7839eb60ff6e606c496e1c848a	08/feb/2013 19:23:46	Bonifici Italia e SEPA	25266.8	435b8226966d2fb40d52bafd6aaa8a93	92a91621f34b668401e8c26050f4e0c6	✗	IT	IT,3269	23.4602697196	✓	✗
15	97221	6da1465327246224216c1c929c339c6f	18/feb/2013 15:09:11	Bonifici Italia e SEPA	50000	f2a7341750c1cc6dc8bea45185a7fe26	60414014d030aa24b4cef90c32fac61f	✓	PT	IT,16232	23.2738702047	✓	✓

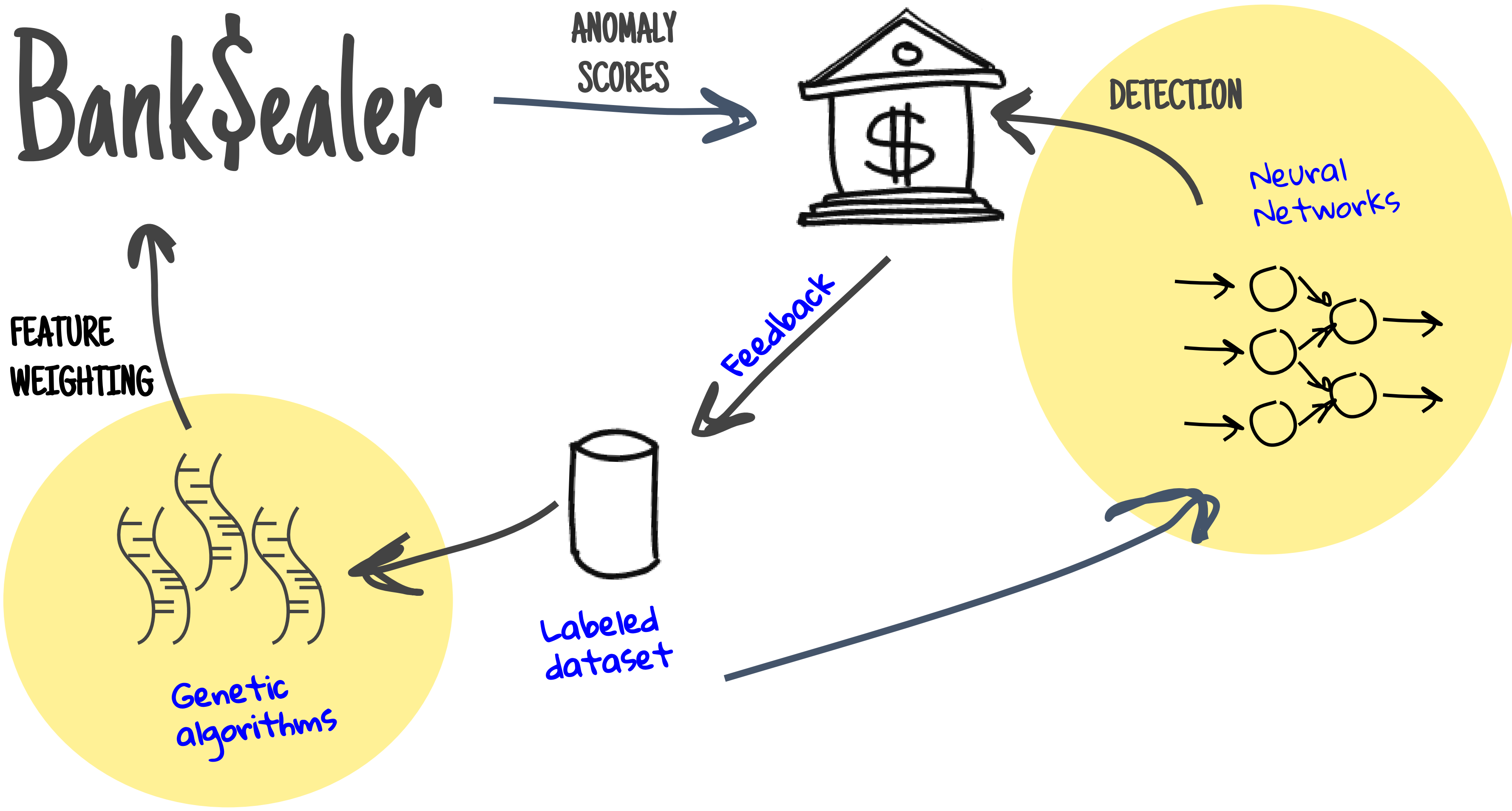
3c9dad3b2601c929e2cece7 ✓ IT IT,30722 24.5660880611 ✓ ✗


BANKSEALER
203 Nuove Transazioni Clicca per ricaricare
Christopher
Rischio | Info | Importo | Utente | IP | IBAN | Operazione | Data
10900€ | Mike Gill | 51.11.1.1 | RO49 AAAA 1B31 0075 93 84 0000 | Bonifico Italia e SEPA | 02/10/2015
Utente poco attivo
560€ | Bruce Carroll | 59.05.21.0 | IT40 3456 5678 4567 34567 0909 987 | Bonifico Italia e SEPA | 02/10/2015
1010€ | Kate Wilkerson | 55.01.09.11 | IT17 X060 5502 1000 0000 1234 567 | Bonifico Italia e SEPA | 02/10/2015
1600€ | Jordan Powers | 151.01.11.02 | DE85 3703 0044 0053 2013 00 | Bonifico Italia e SEPA | 02/10/2015
1030€ | Floyd Houston | 163.01.11.08 | IT40 5054 2811 1010 10000 0123 456 | Giroconto | 02/10/2015
2010€ | Jay Walton | 152.01.11.23 | IT40 5054 2811 1010 10000 0123 456 | Bonifico Italia e SEPA | 02/10/2015
1400€ | Abbie Barnes | 62.01.11.20 | IT40 5054 2811 1010 10000 0123 456 | Bonifico Italia e SEPA | 02/10/2015
1300€ | Mina Harvey | 72.01.11.10 | IT40 5054 2811 1010 10000 0123 456 | Bonifico Italia e SEPA | 02/10/2015
1100€ | Charles Beck | 72.01.11.10 | Filtra per questo elemento | Giroconto | 02/10/2015
90€ | Stanley Morales | 102.01.11.10 | IT17 X060 5502 1000 0000 1234 567 | Bonifico Italia e SEPA | 02/10/2015
1130€ | Antonio Griffith | 102.01.11.10 | IT40 3456 5678 4567 34567 0909 987 | Bonifico Italia e SEPA | 02/10/2015
140€ | Jordan Powers | 101.01.11.10 | IT40 3456 5678 4567 34567 0909 987 | Bonifico Italia e SEPA | 02/10/2015
110€ | William Peters | 14.01.11.10 | IT40 5054 2811 1010 10000 0123 456 | Bonifico Italia e SEPA | 02/10/2015
34€ | Iva Rodgers | 103.02.11.10 | DE85 3703 0044 0053 2013 00 | Bonifico Italia e SEPA | 02/10/2015
21€ | Hester Taylor | 65.01.11.21 | IT40 3456 5678 4567 34567 0909 987 | Bonifico Italia e SEPA | 02/10/2015
Nuovo utente
678€ | Jorge Hopkins | 58.01.12.1 | DE85 3703 0044 0053 2013 00 | Bonifico Italia e SEPA | 02/10/2015
43€ | Emilie Erickson | 14.01.11.10 | IT17 X060 5502 1000 0000 1234 567 | Bonifico Italia e SEPA | 02/10/2015
789€ | Albert Jenkins | 62.01.11.20 | IT17 X060 5502 1000 0000 1234 567 | Bonifico Italia e SEPA | 02/10/2015
56€ | Jesus Reeves | 72.01.11.10 | DE85 3703 0044 0053 2013 00 | Bonifico Italia e SEPA | 02/10/2015
98€ | Beulah Brady | 152.01.11.23 | IT17 X060 5502 1000 0000 1234 567 | Bonifico Italia e SEPA | 02/10/2015
2356€ | Max Kim | 163.01.11.08 | IT17 X060 5502 1000 0000 1234 567 | Bonifico Italia e SEPA | 02/10/2015
Jordan Powers
N° medio transazioni mensili 37
Attivo da Gen. 2014
Spesa media mensile € 3670
Transazione #13412342
Rischio
Rischio medio
Motivazione
Importo 60%
IP 20%
IBAN 10%
Data 4%
Riceve SMS 3%
ASN 2%
IBAN_CC 1%
Segna come interessante
Segnala Frode

Fraud Analysis

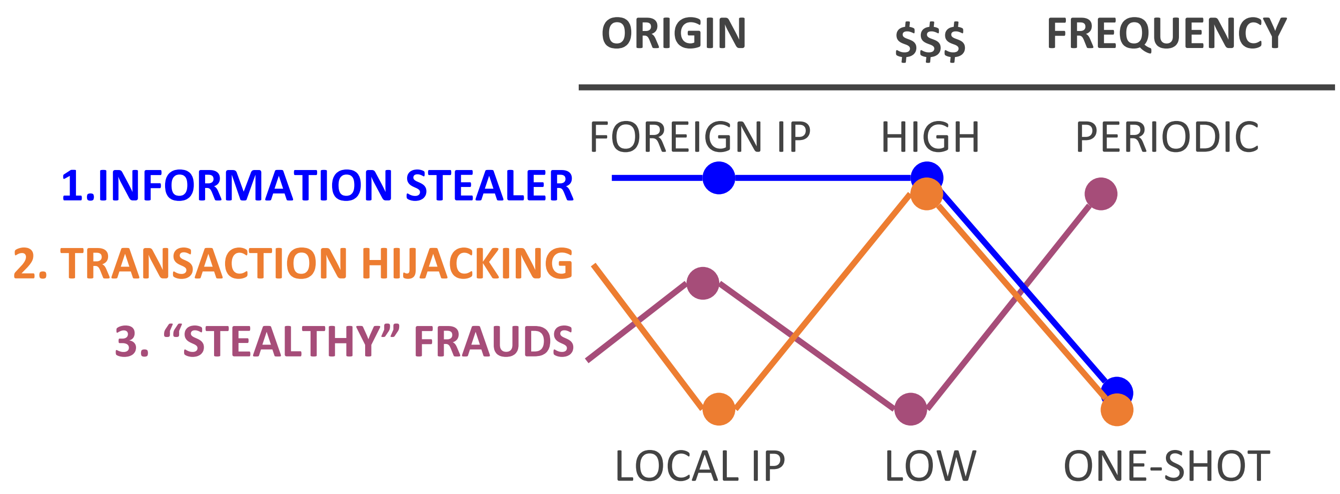

BANKSEALER
203 Nuovi Risultati Clicca per ricaricare
Christopher
1900€ Mike Gill 192.01.09.11 IT40 5054 2811 1010 10000 0123 456 Bonifico 02/10/2015 4:50 pm
130€ Bruce Carroll 192.01.09.11 IT40 3456 5678 4567 34567 0909 987 Operazione 02/10/2015 4:44 pm
789€ Albert Jenkins 192.01.09.11 IT40 5054 2811 1010 10000 0123 456 Operazione 02/10/2015 3:11 pm
56€ Jesus Reeves 192.01.09.11 IT40 5054 2811 1010 10000 0123 456 Operazione 02/10/2015 3:08 pm
98€ Beulah Brady 192.01.09.11 IT40 5054 2811 1010 10000 0123 456 Operazione 02/10/2015 3:02 pm
2356€ Max Kim 192.01.09.11 IT40 5054 2811 1010 10000 0123 456 Operazione 02/10/2015 3:01 pm
Jordan Powers
37
Gen. 2014
€ 3670
Analisi Importo
Distribuzione importo utente
Distribuzione importo utente
COMPORTAMENTO UTENTE
STORICO UTENTE
80
60
40
20
0
42
0-50
50-100
100-200
200-400
400-800
800-1600
1600-3200
3200-6400
12k-20k
20k-40k
40k-50k
Utente: Jordan Powers
Ultimo aggiornamento: 10:45 AM
Segna come interessante
Segnala Frode

Feature Weighting & Detection



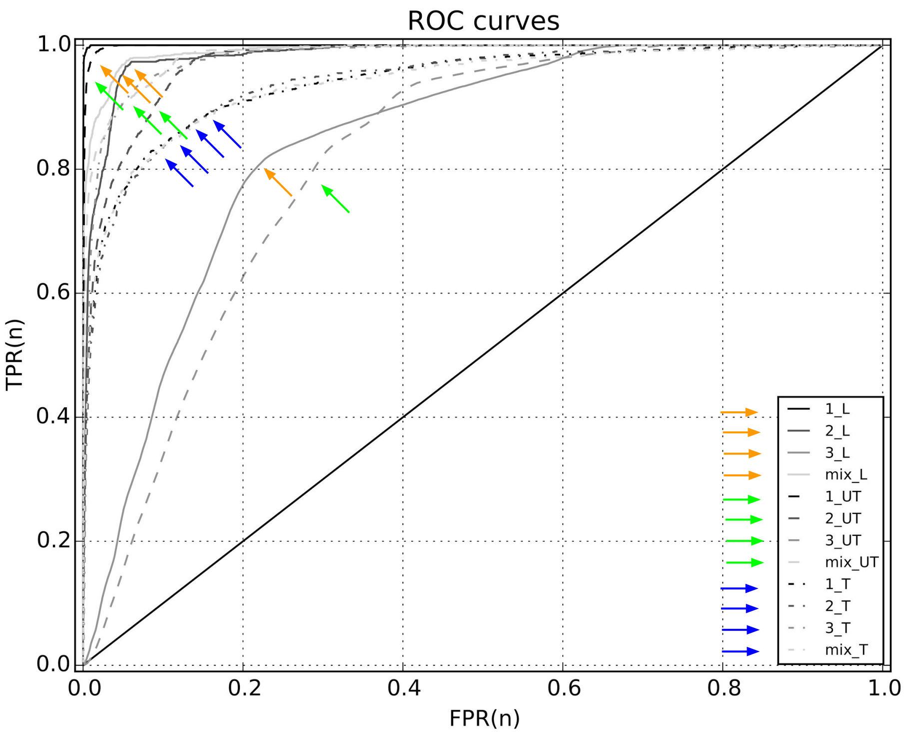
Dataset Generation

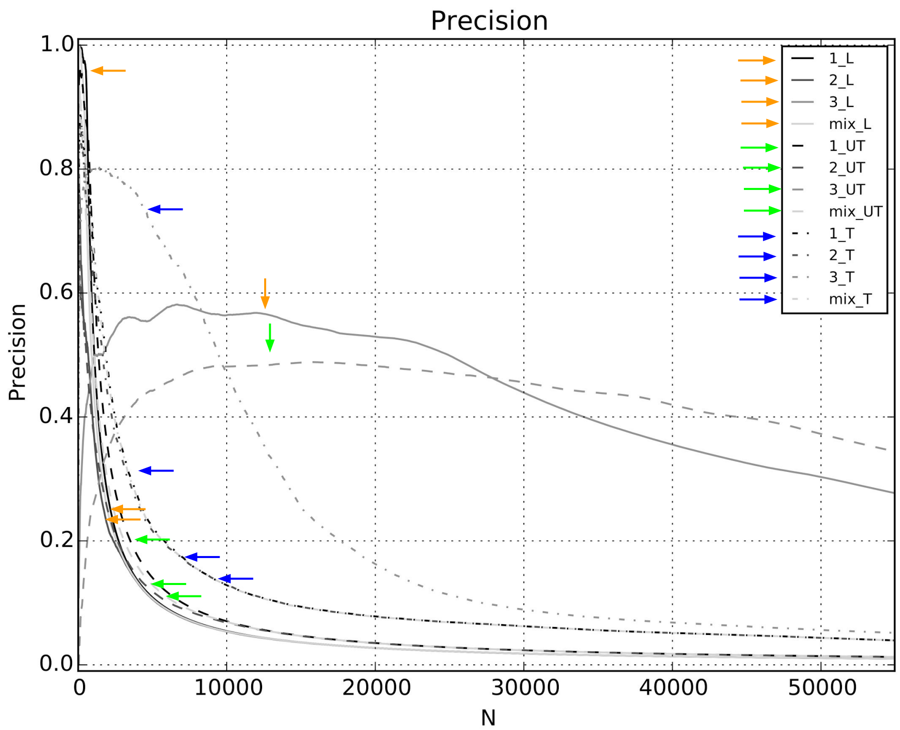
Generate synthetic frauds based on scenarios **built with the collaboration of bank experts** that replicate the **typical real attacks** performed against online banking users



Inject **n fraudulent transactions (or users)** in the testing dataset and analyze the top **n transactions (or users)** in the ranking

Detection Capabilities

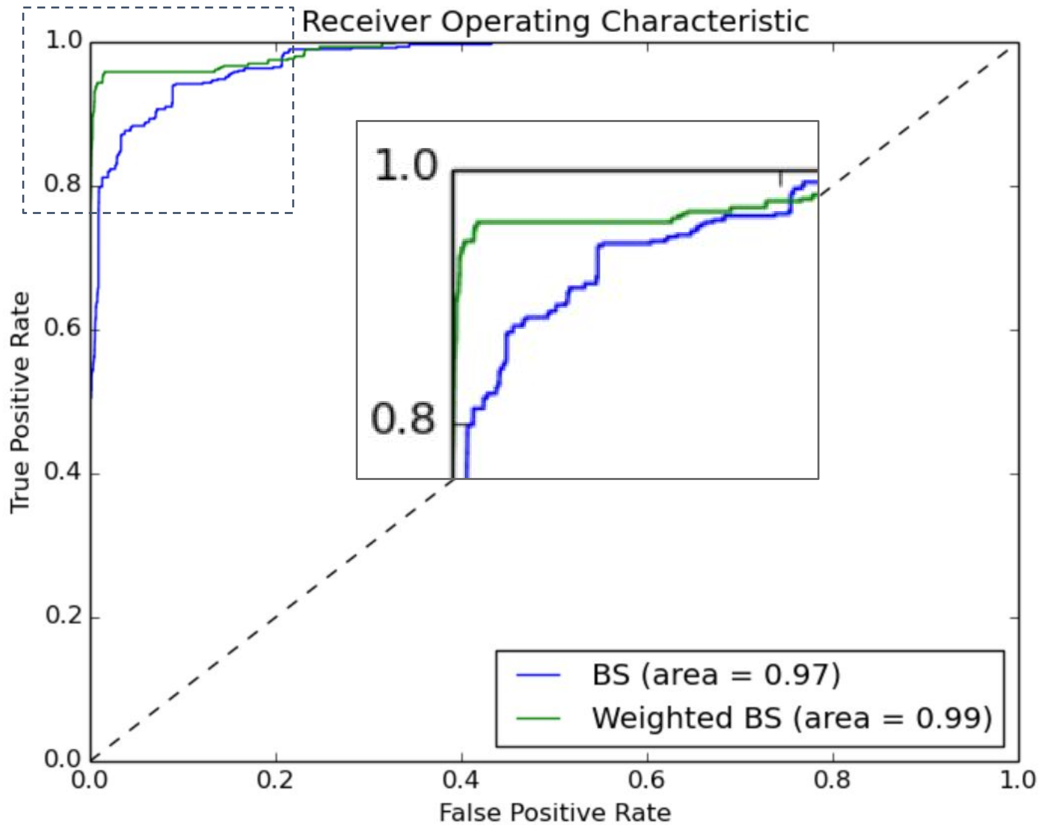




For comparison, best result in **state of the art**: Wei et al. (2013) report detecting **60-70%** of the frauds with unreported precision

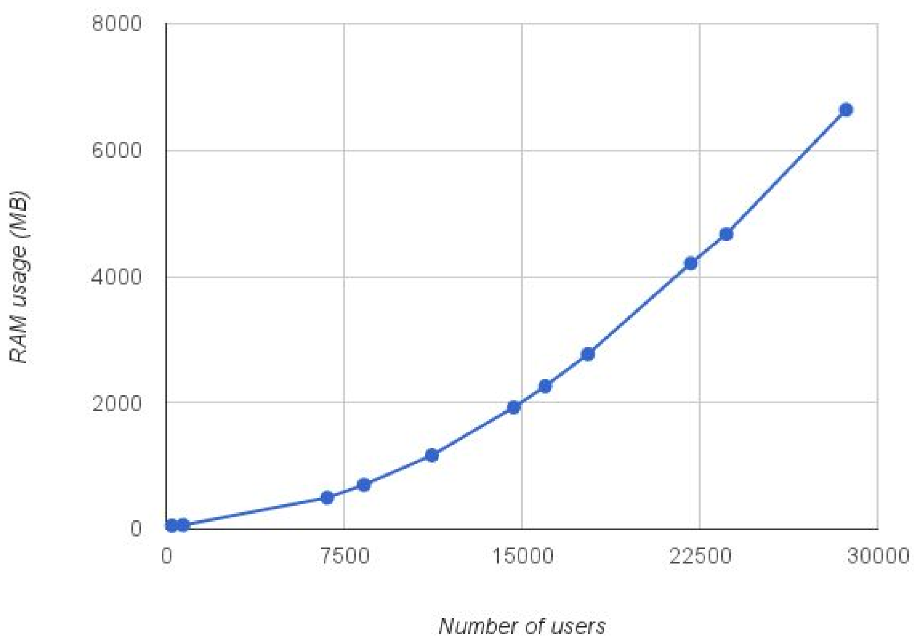
With Genetic Algorithms

	TPR	Weighted TPR	Improvement
Mixed scenario	58%	81%	+23%

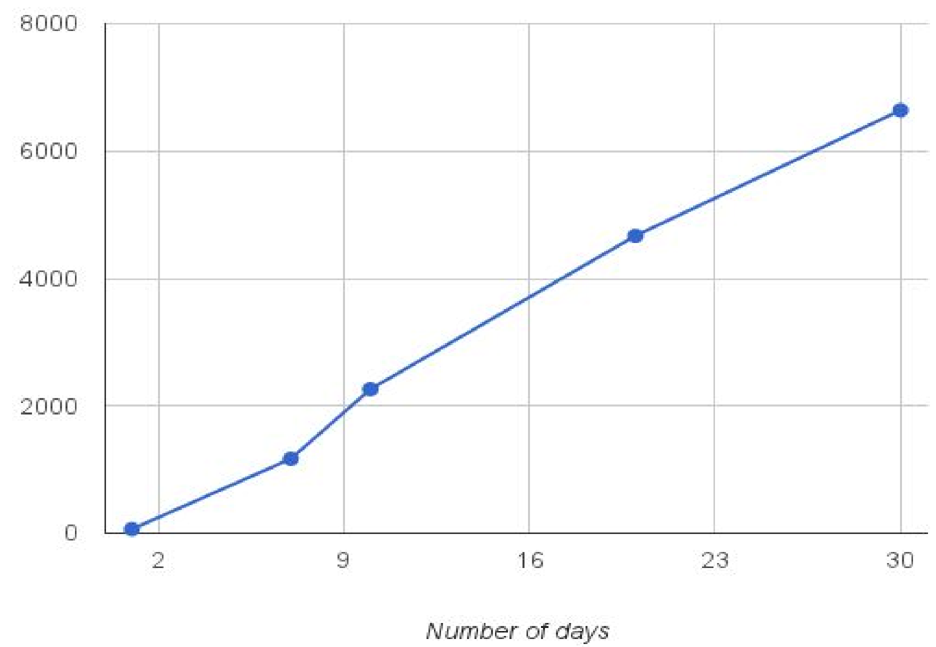


Resource Requirements: Training

RAM per #Users



RAM per #days

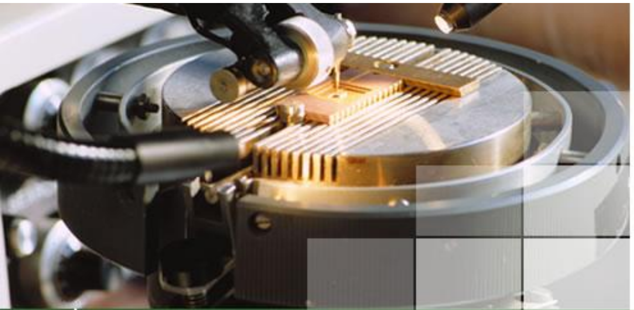


Time Performance: Runtime

	# Transactions	# Users
Bank Transfers	371,137	47,650
Prepaid phone	54,141	16,093
Debit cards	34,986	8,415

Domain	Timespan of Data	Runtime
Bank Transfers	1 day 1 month	1–4 min 6–93 min
Prepaid phone	1 day 1 month	18–25 sec 0.5–2.5 min
Debit cards	1 day 1 month	7–10 sec 12–60 sec

Note: ranges are for "only well trained" and "including undertrained" users.



POLITECNICO
MILANO 1863

Dipartimento di Elettronica
Informazione e Bioingegneria

POLITECNICO DI MILANO

BankSealer: Fast and Transparent Online Banking Fraud Detection and Investigation

Federico Maggi - federico.maggi@polimi.it

Joint work with: Michele Carminati, Stefano Zanero, Ilenia Epifani

NECST laboratory