

Méthodes de détection de communautés pour la détection de fraudes

Safa EL AYEB



Introduction



L'analyse des réseaux sociaux vise à explorer à la fois les structures sociales sous-jacentes et l'influence de chaque caractéristique sur les différentes associations

[Wasserman, 1994]. ”

Objectif



- ✓ L'analyse du réseau social des transactions d'Orange Money.
- ✓ Application des techniques des réseaux sociaux sur des cas d'usage : détection de fraude.

Plan



1. Contexte
2. Axes de recherche
3. Approches existantes
4. Méthodologie
5. Description des données
6. Méthodes d'évaluation



Contexte



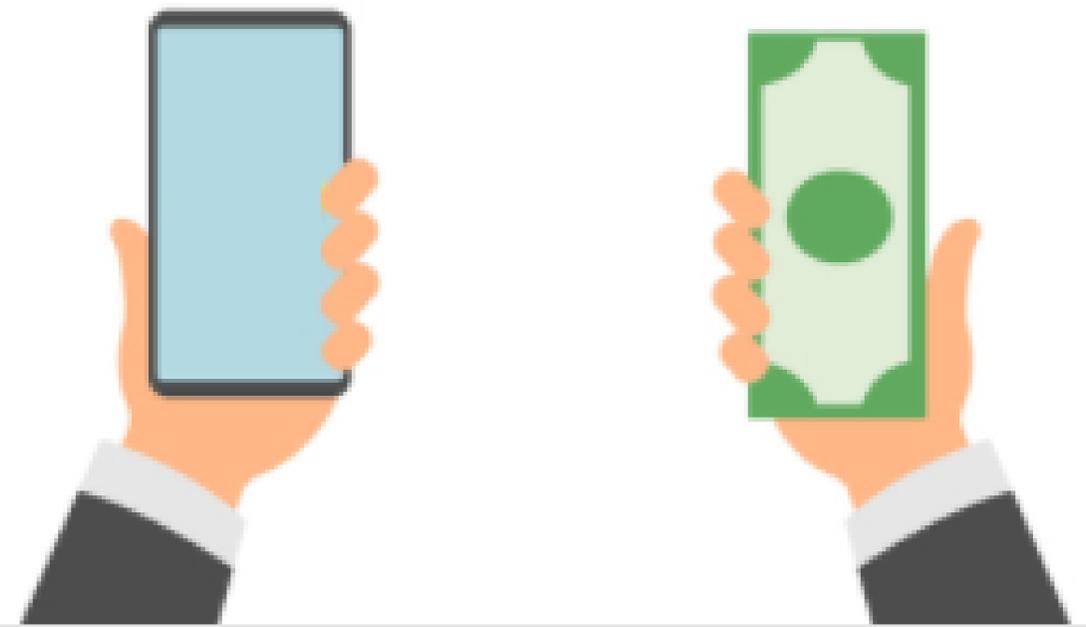
✓ Le Mobile Money ?

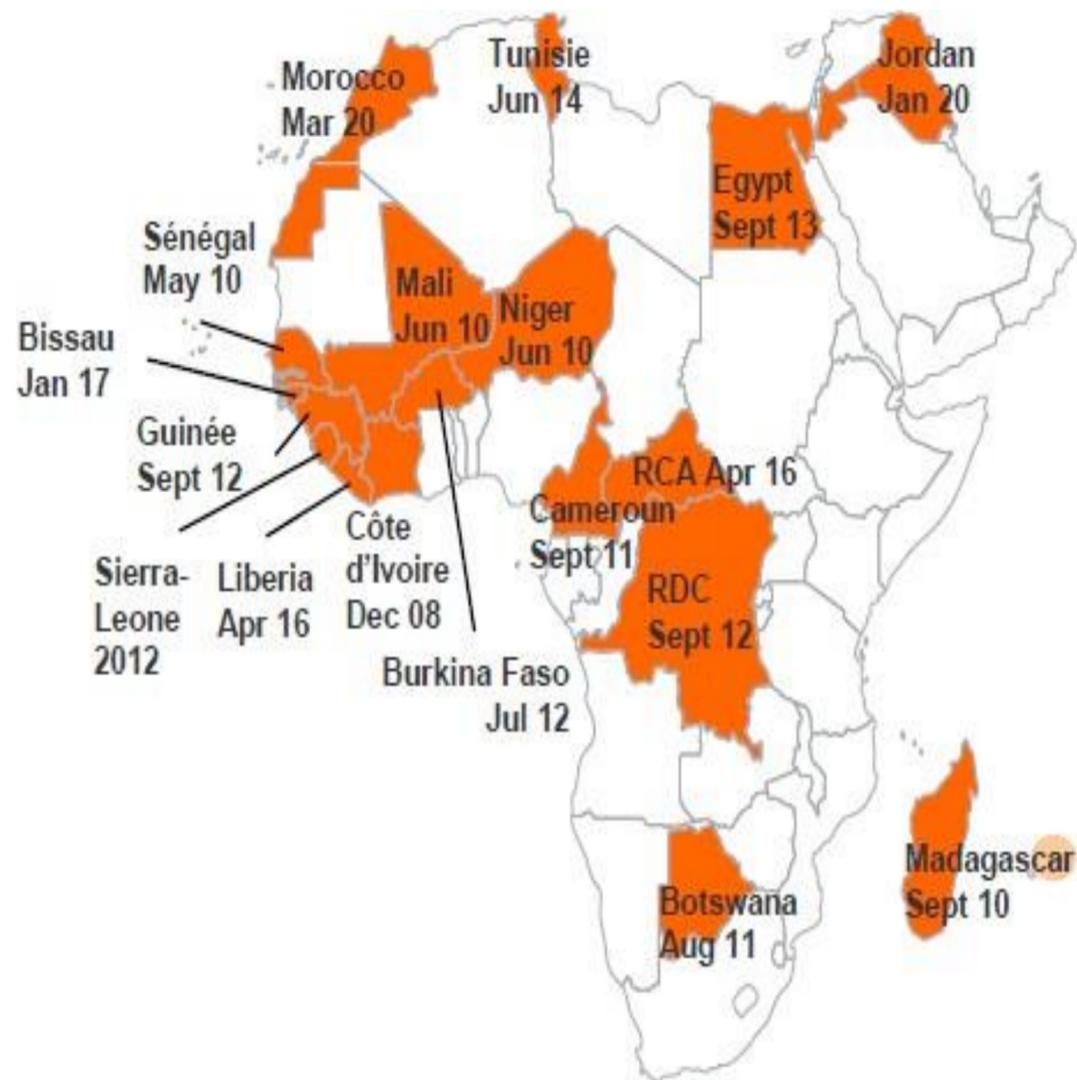
Une alternative au cash et aux systèmes bancaires traditionnels.

Permet de recevoir, conserver et dépenser de l'argent en utilisant le téléphone portable.

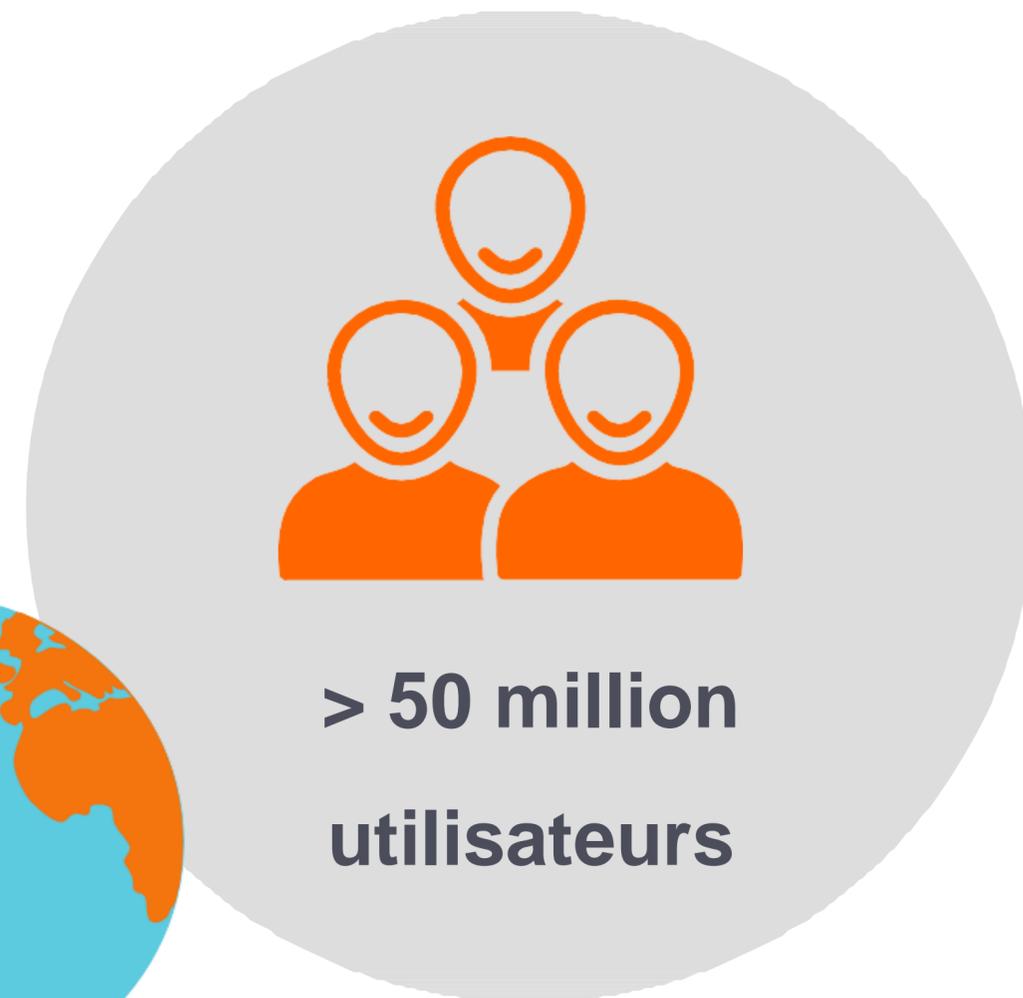
✓ Quels sont ses bénéfices

✓ Comment ça fonctionne ?



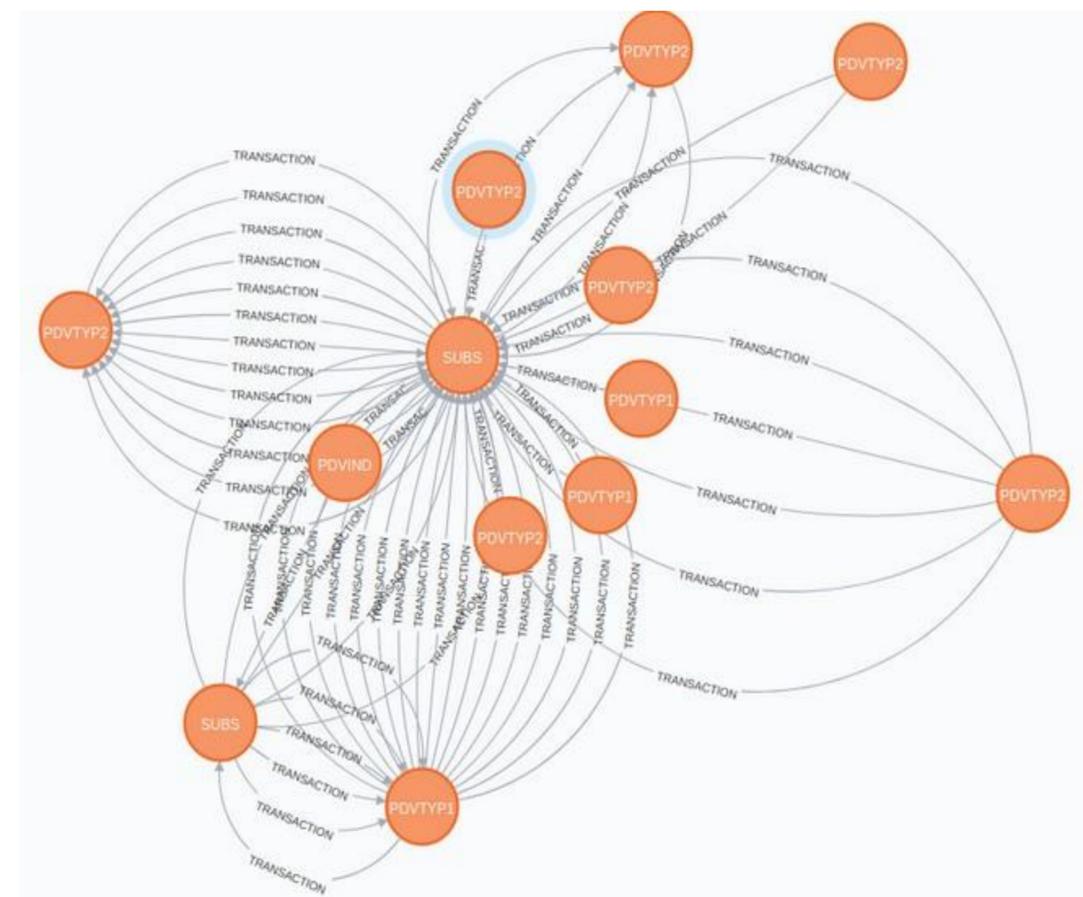


18 pays



SENDER_USER_ID	RECEIVER_USER_ID	SENDER_CATEGORY_CODE	RECEIVER_CATEGORY_CODE	TRANSACTION_AMOUNT	SERVICE_TYPE	TRANSFER_STATUS	TRANSFER_DATETIME
User_i	User_j	Category 1	Category 3	1000	P2P	TS	01/01/2020 9:00:00

Utilisateurs	Transactions
Utilisateur 1	T1
Utilisateur 2	T2
Utilisateur 3	T3
....
Utilisateur n	Tn



3 Millions d'utilisateurs

90 Millions de transactions

Axes de recherche



01

Analyse des données
transactionnelles

02

Utilisation des bases de
données orientées graphe

03

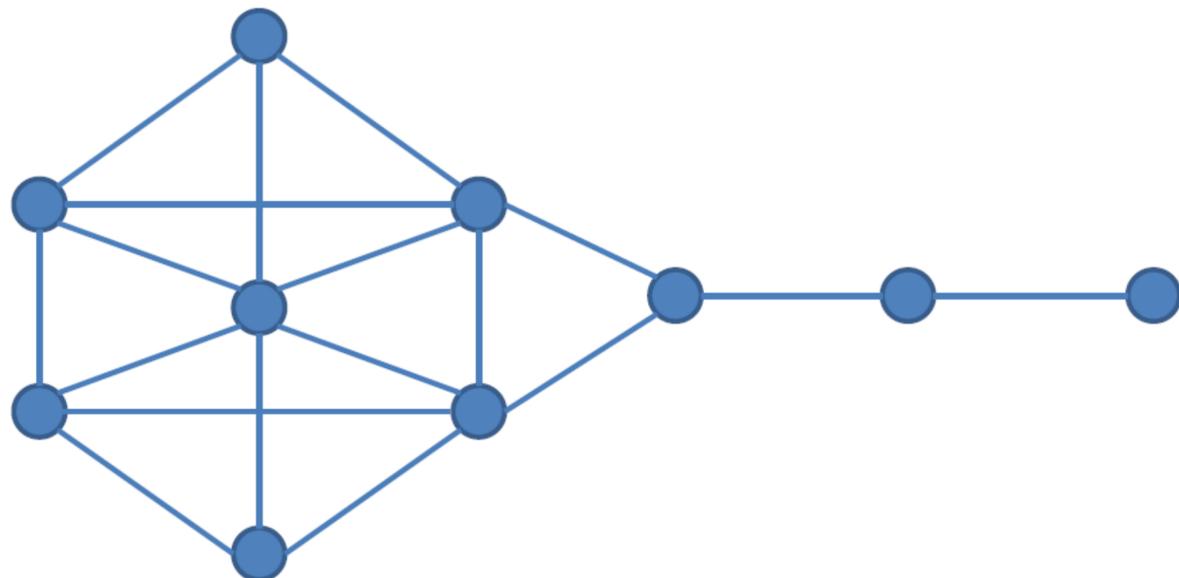
Détection des communautés

Propriétés des graphes



Un graphe se présente sous la forme :

$$G = (V, E)$$



- ★ Diamètre
- ★ Densité
- ★ Modularité
- ★ Centralité
- ★ Clustering coefficient

Méthodes de détection de communautés



Méthodes spectrales

Méthodes de propagation de label

Méthodes hiérarchiques

Agglomérative
Divisive

Méthodes de détection de communautés

Méthodes basées sur la modularité

Greedy techniques
Simulated annealing
Extremal optimization

Cas d'usage : SMiShing



- ★ Une forme d'escroquerie à travers les SMS.
- ★ Les SMS sont souvent en forme d'alerte : une banque, l'opérateur téléphonique, la loto, etc.



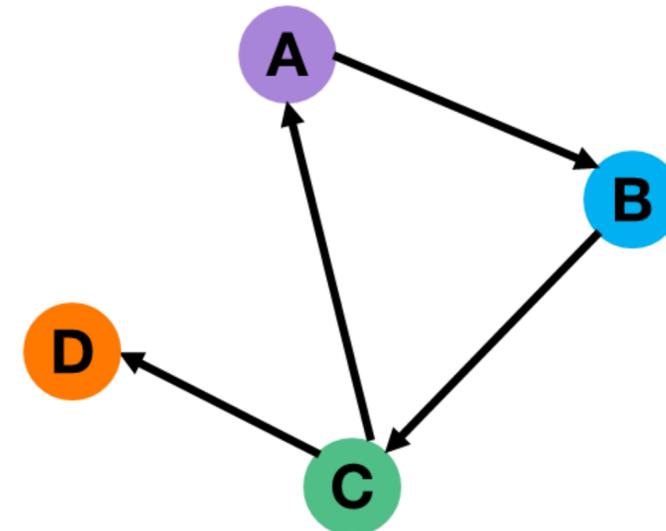
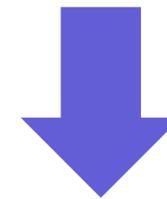
Description des données



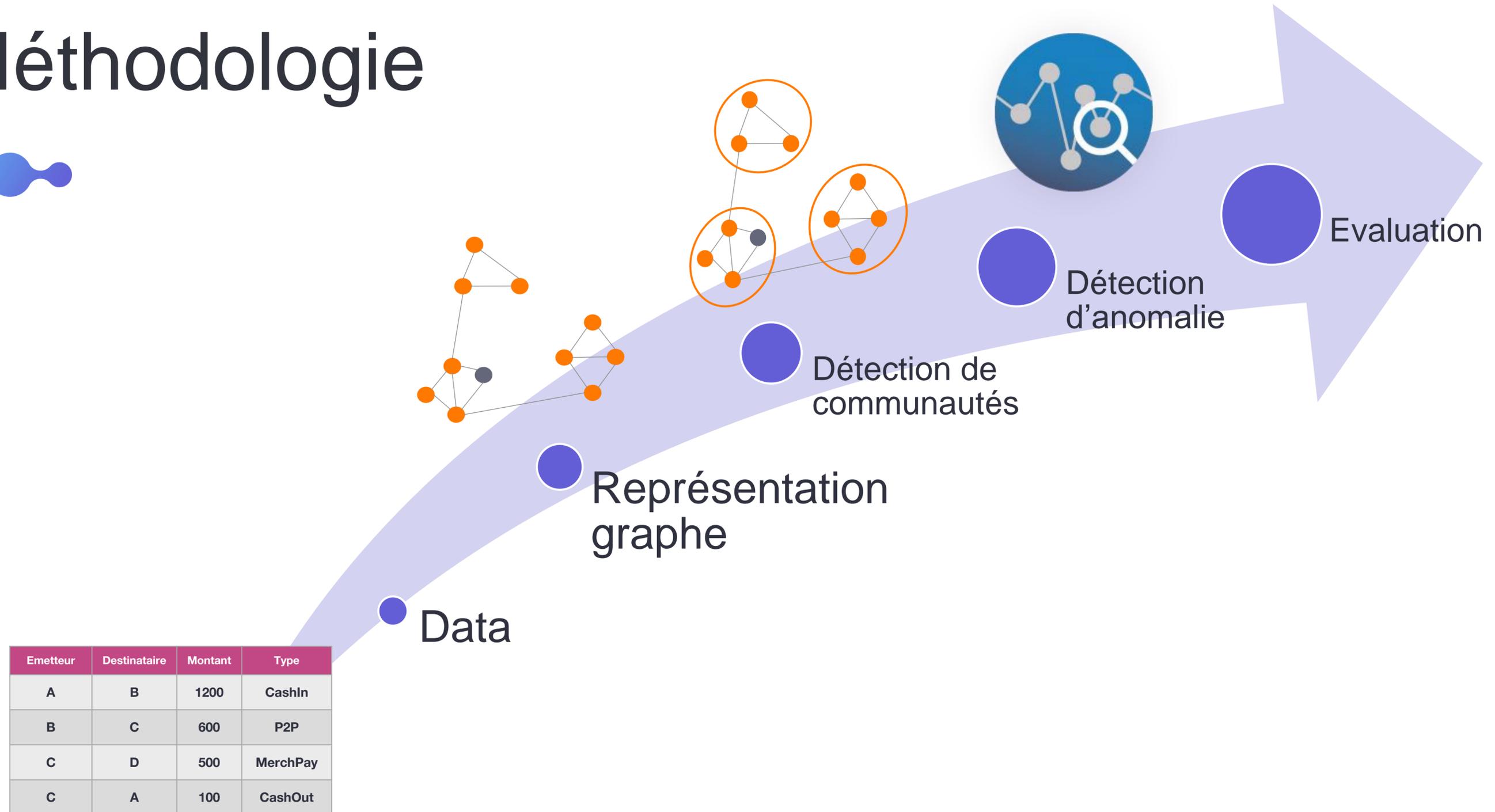
Les données contiennent les transactions entre clients et agents.

4 types de transactions : Cash-In, Cash-Out, merchant payments, and peer-to peer.

Emetteur	Destinataire	Montant	Type	Statut
A	B	1200	CashIn	TS
B	C	600	P2P	TS
C	D	500	MerchPay	TS
C	A	100	CashOut	TF



Méthodologie

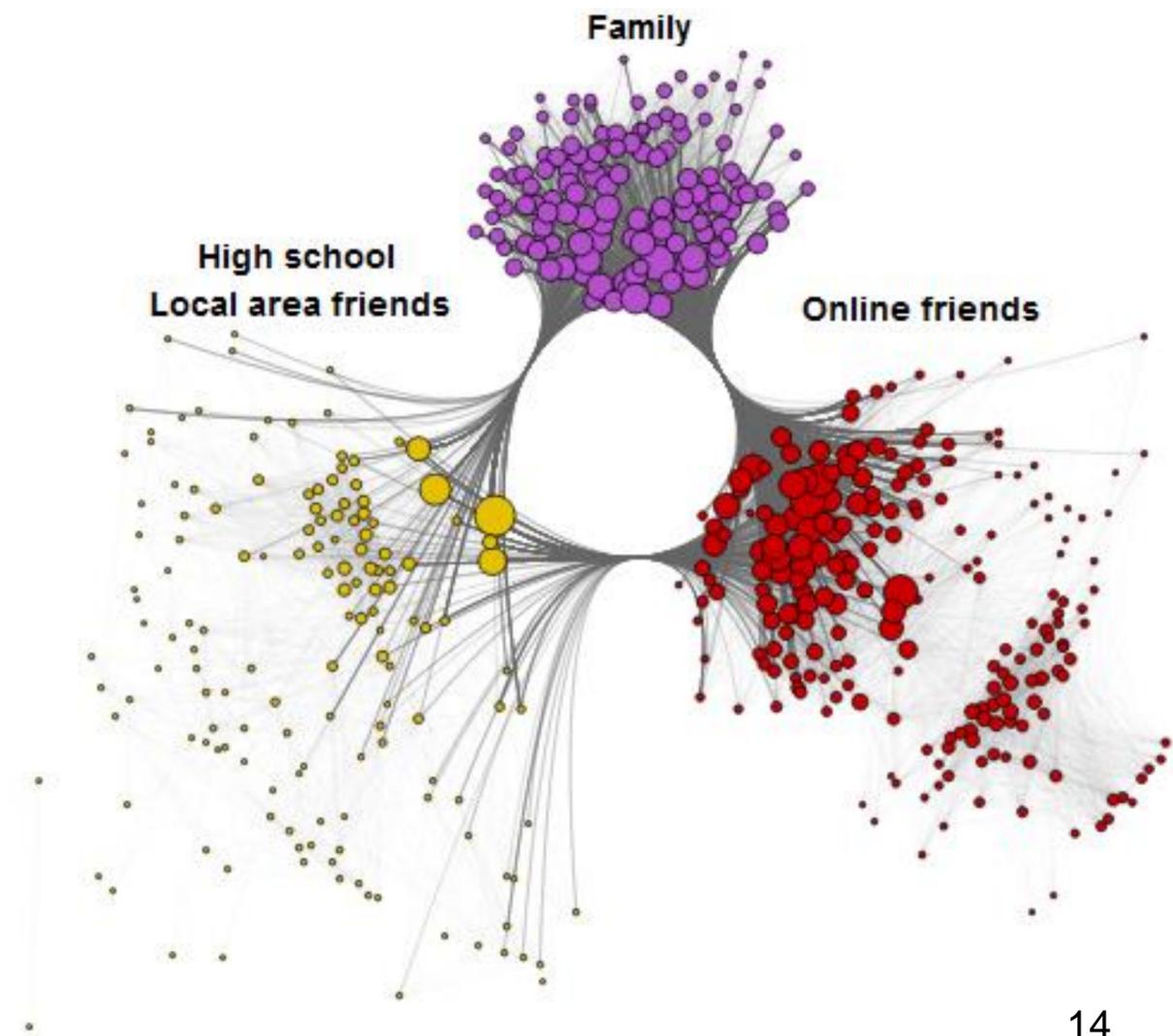


Emetteur	Destinataire	Montant	Type
A	B	1200	CashIn
B	C	600	P2P
C	D	500	MerchPay
C	A	100	CashOut

Evaluation



- ✓ Les algorithmes de détection de communautés explorés dans l'état de l'art seront appliqués
- ✓ Pour évaluer les algorithmes :
 - Métriques existantes : Modularité, NMI, F1-score, etc.
 - Construire une nouvelle métrique.

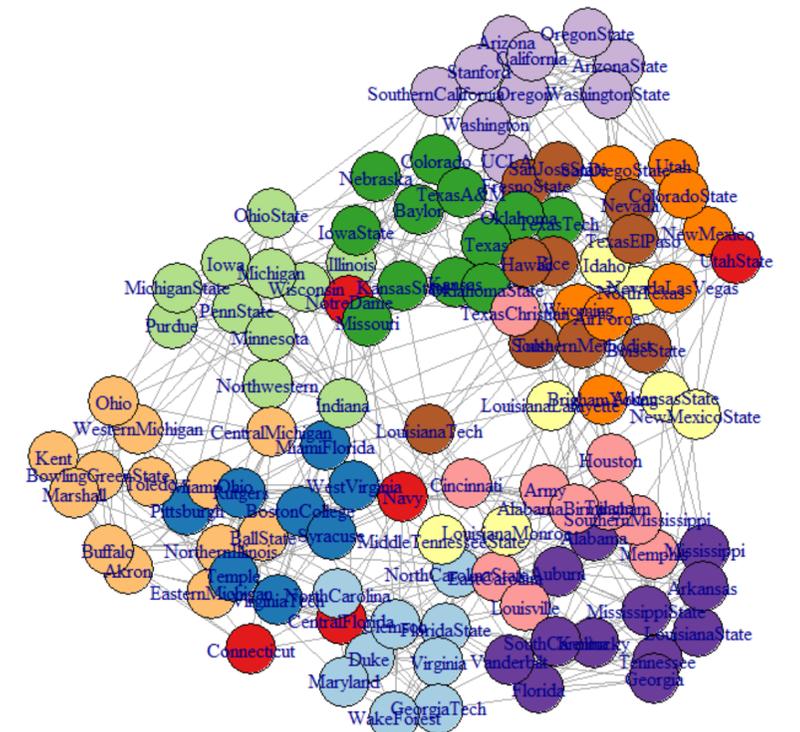
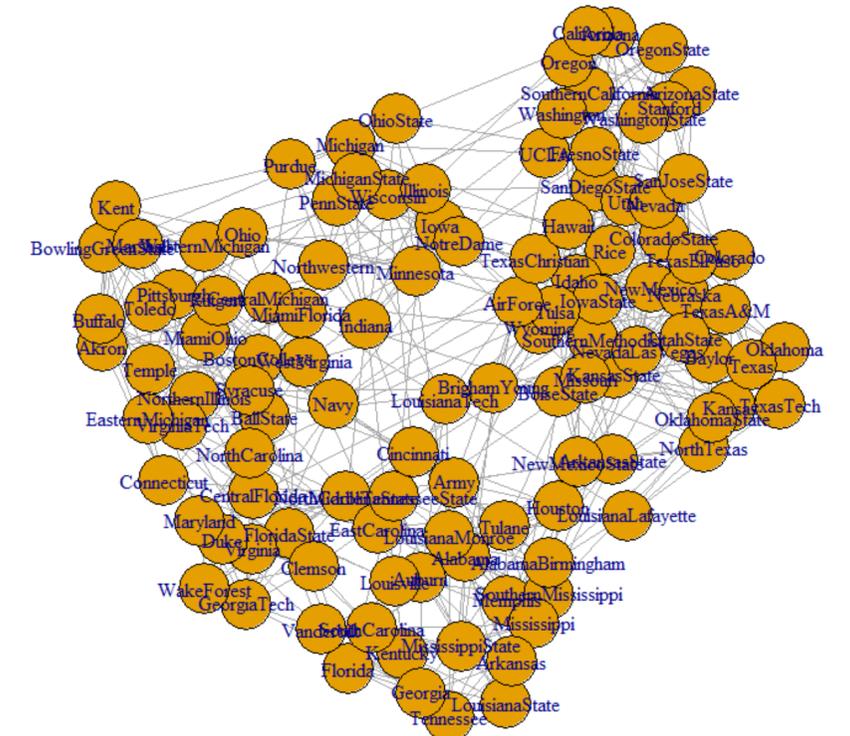


Bases de données de tests



Réseau	Zachary karate club	Football College	Dolphins	EU-Email core
Nombre des nœuds	34	115	62	1005
Nombre des arêtes	78	613	159	25571
Nombre de communautés	2	12	2	42

Tests

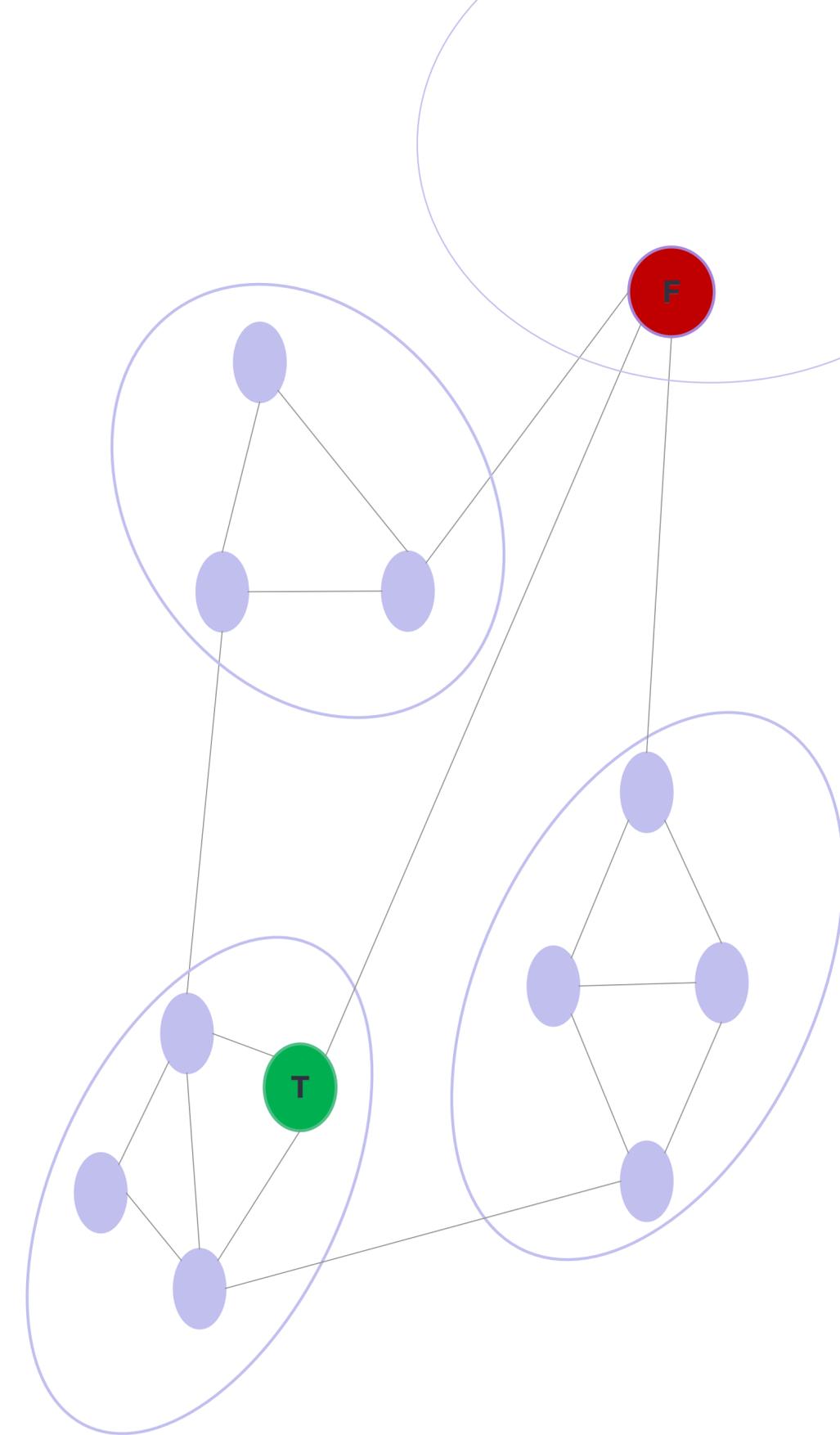


	Degré interne	Densité interne	Modularité	Taille	Conductance	NMI	F1
Lp	7.130	0.203	0.583	10.454	0.334	0.920	0.911
Louvain	7.701	0.182	0.604	12.777	0.281	0.951	0.947
GN	8.523	0.110	0.509	28.75	0.200	0.650	0.687
Greedy modularity	7.720	0.120	0.549	19.166	0.277	0.734	0.713
Walktrap	7.571	0.187	0.602	11.5	0.294	0.907	0.884
Leiden	7.572	0.190	0.604	11.5	0.294	0.930	0.885
Spinglass	7.331	0.205	0.601	10.454	0.314	0.935	0.901

Evaluation de la détection de fraude



- ✓ Transactions à l'intérieur des communautés considérées légitimes.
- ✓ Evaluer la détection des transactions frauduleuses : précision, "recall", etc.



Conclusion



Notre étude a abordé le problème de la détection des communautés au sein des réseaux sociaux afin d'identifier les acteurs frauduleux.

Applications : Visualisation, analyse des flux monétaires, détection des leaders.

Perspectives : Analyse des réseaux dynamiques, analyse des multi-graphes et graphes multi-couches.

Merci de votre attention

