Sistemi Peer-To-Peer Analisi della Lightning Network

Carlo Cantamaglia 0000895868

SOMMARIO	
Introduzione	
Strumenti	1
Metodi	2
Overall metrics	2
Vertices metrics	3
Ego Network	3
Analisi	5
Dataset	5
Overall metrics	6
Vertices metrics	9
Risultati	14
Criticità della rete	17
Topology-Based Attacks	19
DoS	19
Esaurimento dei canali & Isolamento nodi	20
Conclusioni	23
Sitografia	24

Introduzione

Il concetto di Lightning Network (LN) è stato creato da Joseph Poon e Thaddeus Dryja nel 2015.

L'idea principale alla base del progetto è l'elaborazione di un protocollo di pagamento che possa essere usato come soluzione *off-chain* per aumentare scalabilità, usabilità, velocità e diminuire le spese di commissione di Bitcoin.

Il LN è composto da un network di trasferimento *off-chain* sviluppato sulla *blockchain* di Bitcoin. Il sistema opera a livello *peer-to-peer* (P2P) e il suo utilizzo si basa sulla creazione dei cosiddetti canali di pagamento bidirezionali, attraverso i quali gli utenti possono eseguire transazioni dirette di criptovalute.

Nel network, data la capacità di instradare i pagamenti, non sono presenti canali tra tutti i nodi, ciò potrebbe addirittura essere visto come uno spreco di risorse.

In questo progetto è stata eseguita un'analisi approfondita sulla LN per vederne le caratteristiche e gli eventuali punti di debolezza, confrontando i risultati con alcune delle statistiche presenti in rete.

Strumenti

NodeXL

NodeXL è un pacchetto *software* di analisi e visualizzazione delle reti per *Microsoft Excel*. Nell'analisi è stato utilizzato per il calcolo di alcune delle statistiche, come ad esempio quelle relative alle metriche generali o ai singoli attori, e per la parte grafica.

Ucinet

Ucinet è un pacchetto *software* per l'analisi di reti sociali ma ha un numero maggiore di procedure utilizzabili nell'analisi.

Nel progetto è stato utilizzato principalmente per l'analisi dell'ego network.

NetworkX

NetworkX è pacchetto Python per la creazione, la manipolazione e lo studio della struttura, delle dinamiche e delle funzioni di reti complesse.

Nell'analisi è stato utilizzato principalmente per la verifica e i grafici relativi alla *Small-World* network e *Free Scale network*.

Metodi

L'analisi è stata effettuata analizzando i risultati a diversi livelli. Di seguito verranno spiegate teoricamente le principali misure utilizzate.

Overall metrics

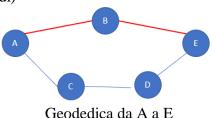
Tramite le metriche generali è stata fatta una prima analisi dell'intera rete senza soffermarmi sui ruoli che ricoprono i singoli all'interno di quest'ultima.

Le principali misure calcolate sono:

- Nodes: numero di attori presenti nella rete
- Edges: numero di legami tra i vari attori
- Edges with duplicate: numero di tutti i legami presenti più di una volta tra due attori
- Components: numero di sottografi connessi nel grafo principale



• Average geodesic distance: media tra tutte le geodediche all'interno del grafo (geodedica: percorso più breve tra due nodi)



- Diameter: lunghezza della più grande distanza geodetica all' interno del grafo
- *Density*: proporzione tra il numero di legami che sono attualmente presenti nel grafo e il numero di legami che potrebbero essere presenti:

$$\Delta = \frac{L}{g(g-1)/2} = \frac{2L}{g(g-1)}$$

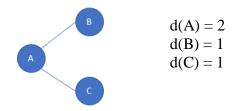
- Average Degree: media tra tutti i gradi degli attori presenti nella rete (Degree: spiegato nella sezione "Vertices metrics")
- Average Betweeness Centrality: media tra tutte le betweeness centrality degli attori nella rete (Betweeness centrality: spiegata nella sezione "Vertices metrics")

Vertices metrics

Tramite l'analisi delle metriche relative ai singoli vertici sono stati individuati quali ruoli ricoprono all' interno della rete.

Le principali misure calcolate sono:

• Degree: numero di legami che collegano l'attore agli altri.



• Betweeness Centrality: misura della centralità di un attore in un grafo basato sui percorsi più brevi, ovvero il numero di geodediche che attraversano il vertice.

Ego Network

Tramite l'analisi delle *Ego Network* è stata descritta e modellizzata la variazione tra gli individui nel modo in cui sono interlacciati nelle strutture sociali "locali".

Le *Ego Network* sono costituite da un nodo focale ("ego") e dai nodi a cui l'ego è direttamente connesso ("alter") più i relativi legami.

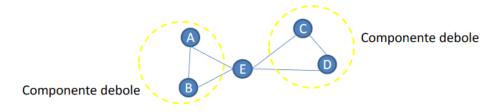
L'insieme dell'ego, degli alter e di tutti i legami formano il cosidetto vicinato.

Nel nostro caso andremo ad analizzare il vicinato "one step" che include solo gli alter che sono direttamente collegati con l'ego.

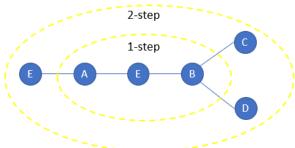
Le principali misure calcolate sono:

- Size: dimensione della rete dell'Ego, ovvero il numero di nodi raggiunti da quest'ultimo.
- Ties: numero di connessioni dirette tra tutti i nodi all'interno dell'ego network
- Pairs: numero di coppie ordinate, ossia il numero di possibili legami orientati
- Density: numero di legami fratto il numero di coppie
- *nWeakComp*: maggior numero di attori che sono collegati trascurando la direzione dei legami.

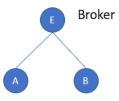
Se l'ego fosse collegato ad A e B (che sono collegati tra loro), e l'ego fosse collegato a C e D (che sono collegati tra loro), ma A e B non sono collegati in alcun modo a C e D (tranne che per il fatto che tutti sono collegati a lego), allora ci sarebbero due "componenti deboli" nel vicinato dell'ego.



- *pWeakCOmp*: Numero di componenti deboli diviso per la dimensione dell'*ego network* (componenti deboli in percentuale)
- *Two-step reach*: numero degli attori della rete che rientrano nei «due passaggi diretti» dall'ego



- Reach efficiency: rapporto tra il numero degli attori raggiungibili tramite due step fratto la dimensione dell'intera rete.
- *Brokerage*: numero di coppie non direttamente connesse. L'idea del *Brokerage* è che l'ego è l'intermediario per le coppie di altri attori. In una rete di Ego, l'ego è collegato a tutti gli altri attori (per definizione). Se questi altri non sono collegati direttamente tra loro, l'ego può essere un broker. Un elemento di interesse è semplicemente quanto potenziale di intermediazione esiste per ciascun attore (quante volte coppie di vicini nella rete dell'ego non sono direttamente collegate).



• *nBrokerage*: broker normalizzato, ossia il brokerage diviso per numero di coppie. Questo numero è un indicatore di quanto l'Ego è un intermediario importante. Un nodo può trovarsi in una posizione di intermediazione diverse volte, ma questa potrebbe essere una piccola percentuale del totale delle connessioni possibili in una rete (ad esempio se la rete è grande).

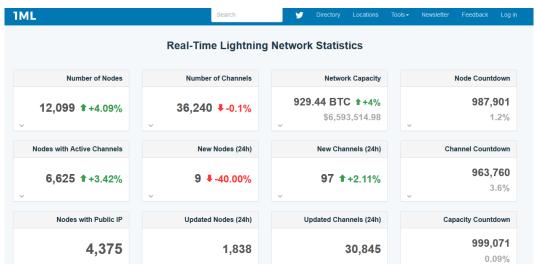
Analisi

Dataset

Il database selezionato comprendeva sia canali chiusi che aperti, ma per la seguente analisi sono stati selezionati solamente quelli ancora aperti.

Il grafo derivante può essere visto come un grafo non-direzionale one-mode in quanto ogni canale di pagamento è bidirezionale e l'analisi verrà effettuata solamente su un tipo di legame.

In totale sono presenti nel dataset 5910 vertici e 34281 *edges*, dati abbastanza veritieri in confronto alla grandezza attuale della LN che come possiamo notare ha subito un aumento nell'ultimo periodo: 6625 nodi con canali attivi e 36240 canali attivi.



1ML.com-Ral Time LN statistic

Graph Metric	Value 🔻
Graph Type	Undirected
Vertices	5910
Unique Edges	28291
Edges With Duplicates	5990
Total Edges	34281
Self-Loops	0
Connected Components	32
Single-Vertex Connected Components	0
Maximum Vertices in a Connected Component	5843
Maximum Edges in a Connected Component	34244

NodeXL- Overall Metrics

Nel pre-processamento dei dati sono stati effettuati due passaggi:

• Selezione del componente principale:

Il grafo è formato da 32 componenti, il più grande dei quali ha 5843 vertici.

Essendo in totale 5910 vertici ciò significa che i restanti 67 formano 31 componenti, ognuno dei quali composto solamente da due o tre nodi.

Per questo motivo, dopo aver constatato che le statistiche globali restano sostanzialmente invariati, sono stati eliminati dall'analisi, concentrandosi, quindi, sul componente principale.

• Eliminazione degli *edges* con i duplicati:

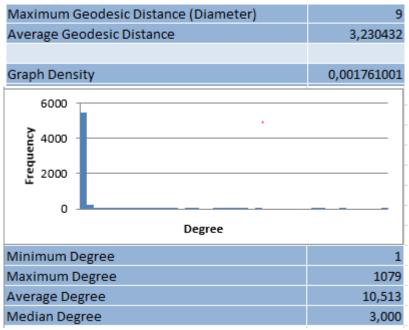
Gli "Edges With Duplicates" nel nostro network implicano la presenza di due nodi con più di un canale tra di loro.

Questi canali aggiuntivi vengono creati per aumentare i "limiti" disponibili per la ricezione e l'invio di criptovalute traminte LN.

Ciò non va ad influire nella nostra analisi e per questo motivo gli *edes* con duplicati sono stati aggregati in un unico *edge*.

Overall metrics

Come prima sono state analizzate le *overal metrics* della rete.



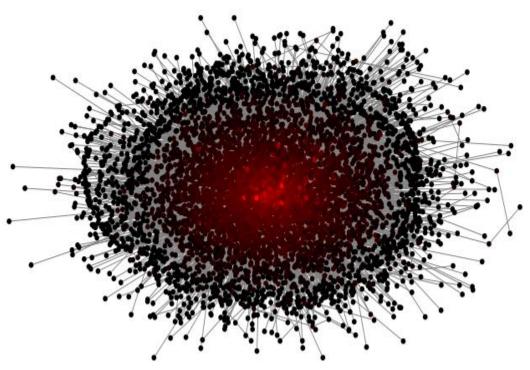
NodeXL-Overall Metrics

Dalla tabella precedente si può notare come la densità è pari allo 0,0017% ciò implica che sono presenti meno del 2% di tutti i legami possibili.

La distanza media tra tutte le geodediche nel *network* è di 3,2 ciò significa che in media due attori all'interno del nostro network utilizzerebbero 3 *channels* per un movimento monetario.

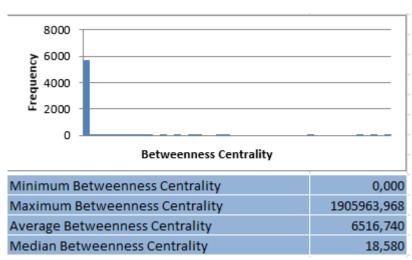
Per quanto riguarda il diametro, ovvero la più grande distanza geodedica, vediamo che nonostante la scarsa densità del grafo è solamente 9 e andando ad analizzare i gradi possiamo notare il grado medio del *network* è 10,5 con un range che parte da 1 ed arriva fino a 1079.

Ciò ci suggerisce una grande variabilità di gradi tra i nodi presenti nella rete che ricoprono quindi ruoli diversi in quest'ultima come vedremo in seguito.

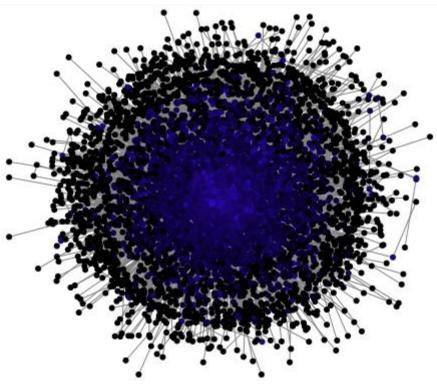


Vertici colorati in base al grado (utilizzando un *logarithmic mapping*)

Questà variabilità, come si può notare, viene confermata anche dalla Betweeness Centrality.



NodeXL-Betweenness Centrality



Vertici colorati in base alla *Betweennes Centrality* (utilizzando un *logarithmic mapping*)

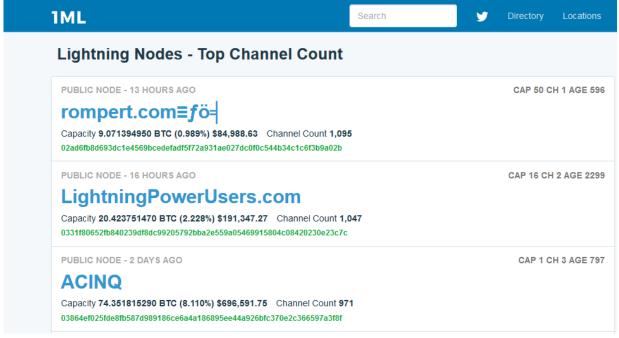
Vertices metrics

Concentrando la nostra analisi sui singoli attori possiamo notare subito che quelli con un elevato grado hanno anche un elevata *Betweenness Centraly*, possiamo quindi classificare questi ultimi con un ruolo di figure centrali.

	Label	`		In-	Out-	Betweenness
Vertex 💌	Position 💌	Tooltip 🕶	Degree 🚚	Degree ▼	Degree 💌	Centrality 🔽
02ad6fb8c			1079			1766326,470
0331f8065			932			1905963,968
03864ef02			862			1859714,260
0217890e3			839			1436977,193
0279c22ec			649			892568,830
03bb88ccc			640			850477,359
0395033b2			591			918145,779
0242a4ae0			571			741323,902
03abf6f44			544			619127,604
03c2abfa9			525			666676,502
0390b5d44			484			589491,636

NodeXL- Vertices metrics

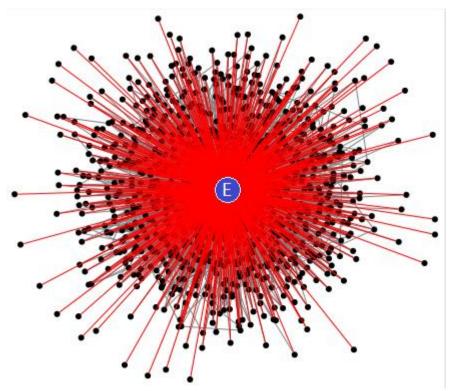
Andando a confrontare i nodi più centrali nella nostra analisi con i *Top Channel* presenti su *1ML* possiamo trovare immediatamente un riscontro, infatti come si può vedere dall'immagine seguente gli attori corrispondono.



1ML.com-LN Top channel

Ego Network

Come ego nell'analisi è stato scelto l'attore 02ad6fb8d693dc1e4569bcedefadf5f72a931ae027dc0f0c544b34c1c6f3b9a02b (rompert.com). L'analisi effettuata su quest'attore può facilmente essere effettuata, con risultati simili, per le altre figure centrali della rete.



Ego Network- Edges dell'ego evidenziati di rosso

La dimensione della rete dell'ego è pari a 1079, ciò trova riscontro con i dati attuali pubblicati su *1ML*.

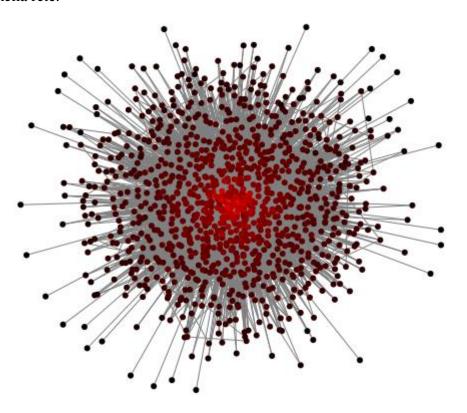
id ▼	Size 🚚	Ties 💌	Pairs 💌	Density 💌
02ad6fb8d	1079	22986	1163162	1,9761649
0331f8065	932	11282	867692	1,300231
03864ef02	862	12294	742182	1,656467
0217890e3	839	8484	703082	1,2066871
0279c22ed	649	16462	420552	3,9143791
03bb88ccc	640	10914	408960	2,6687207
0395033b2	591	10220	348690	2,9309702
0242a4ae0	571	11902	325470	3,6568654
03abf6f44	544	12030	295392	4,0725546
03c2abfa9	525	10154	275100	3,6910214

Ucinet-Ego basic measures

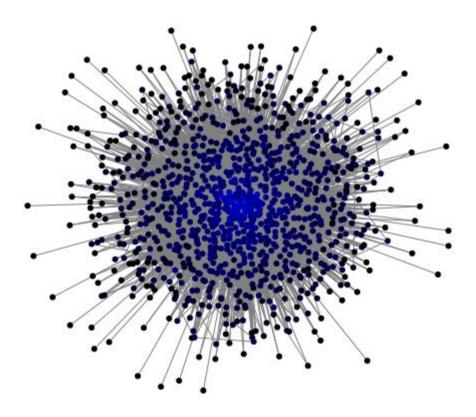


1ML.com-Channel Count Ranking

Tra tutti i nodi dell'ego *network* ci sono 22986 ties con 1.163.162 coppie ordinate, ciò comporta una densità bassa nella rete.



Ego Network - Vertici colorati in base al grado (utilizzando un logarithmic mapping)



Ego Network - Vertici colorati in base alla Betweennes Centrality (utilizzando un logarithmic mapping)

nWeakComp 💌	pWeakComp 🔻	2StepReach 💌	2StepPct 💌	ReachEffic 💌	Broker 💌	nBroker 💌
33	3,05838728	4756	81,41047668	12,56837845	570088	0,9802384
197	21,13733864	4324	74,01574707	16,23732567	428205	0,9869977
151	17,51740074	4538	77,67887878	15,97606087	364944	0,9834353
141	16,80572128	3932	67,30571747	17,01501656	347299	0,9879331
22	3,389830589	4847	82,96816254	14,27982235	202045	0,9608562
47	7,34375	4438	75,96713257	16,3667202	199023	0,9733128
54	9,137055397	4413	75,53919983	16,79543304	169235	0,9706903
27	4,728546619	4550	77,88428497	15,94253635	156784	0,9634314
33	6,066176414	4353	74,51215363	15,23093033	141681	0,9592745

Ucinet-Ego basic measures

L'ego network in questione ha 33 componenti deboli ovvero il 3% dell'intero ego network. In due step l'ego raggiunge 4756 alter ovvero 81% dell'intero network.

Come si può notare dai valori di brokerage e di brokerage normalizzato l'ego ricopre un ruolo di intermediazione molto importante all'interno della rete.

Colonna1 💌	Degree 💌	EffSize →	Efficiency 💌	Constraint 💌	Hierarchy 💌
02ad6fb8d69	1079	1057,697	0,980256677	0,009513786	0,35511279
0331f80652fk	932	919,8948	0,987011671	0,009016425	0,39738312
03864ef025fd	862	847,7378	0,983454525	0,010536918	0,40146887
0217890e3aa	839	828,8879	0,987947524	0,012842298	0,47149932
0279c22ed7a	649	623,6348	0,960916519	0,013132353	0,328275
03bb88ccc44	640	622,9469	0,973354518	0,015016014	0,38193232

Ucinet-Structural Holes

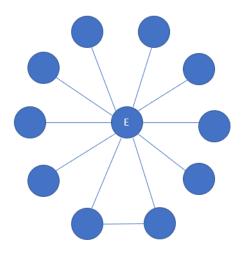
Come abbiamo visto in precedenza la size dell'*ego network* è di 1079, in questo caso possiamo vedere l'effettiva dimensione, ovvero la dimensione totale meno la dimensione di ridondanza totale della rete, che è pari a 1057,697.

Ciò, come indica anche l'efficienza, pari allo 0,98, indica che quasi ogni contatto nella rete non è ridondante.

Per quanto riguarda il vincolo dell'ego possiamo vedere che è molto basso, quasi prossimo allo zero, ciò sta ad indicare che quasi tutte le connessioni dell'ego sono con attori che non sono a loro volta connessi tra di loro.

In fine la gerarchia ci suggerisce che la costrizione dell'ego non è concentrato da una sola relazione (pari a 1) ma non è neanche la stessa per ogni relazione (pari a 0).

Dall'analisi di questa ego network ci aspettiamo quindi una struttura del genere:



Risultati

A seguito dell'analisi effettuata possiamo supporre di caratterizzare la LN come una rete:

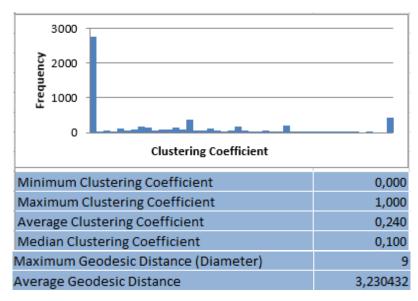
• *Small-world*: Nelle *Small-world network* la distanza tra una qualunque coppia di nodi, definita come il numero di connessioni che li separano lungo il percorso più breve, è molto piccola rispetto al numero dei nodi totale nella rete.

Le reti *Small World* sono caratterizzate da nodi che tendono a raggrupparsi e ad avere un'alta densità di *edges*, più formalmente, il diametro cresce logaritmicamente con il numero di nodi.

Per andare a verificare se la supposizione è corretta ci baseremo su tre misure:

- Average Clustering Coefficent
- Average Geodesic Distance
- Small World Sigma value

Andremo a creare 10 Erdős-Rényi graph in modo randomico da confrontare con la LN.



LN-Clustering & Average Geodesic Distance

Ŧ
37
37
17
16
00
23
21
13
59
03,

Tabella risultati random graph

L'Average Geodedic Distance nella LN, come abbiamo visto, è pari a 3,230443 invece l'Average Clustering Coefficent 0,24.

Per quanto riguarda la media tra i random graph, l'Average Geodesic Distance è uguale a 3,939304 e l'Average Clustering Coefficent 0,0018.

Per calcolare il coefficiente di *Small-world* (sigma) bisogna prima calcolare:

- λ: Average Geodedic Distance della LN / Average Geodedic Distance medio dei random graph
- Υ : Clustering Coefficent della LN / Clustering Coefficent medio dei random graph Dalla nostra analisi λ è uguale a 0,820054 e Υ è uguale a 164,0031.

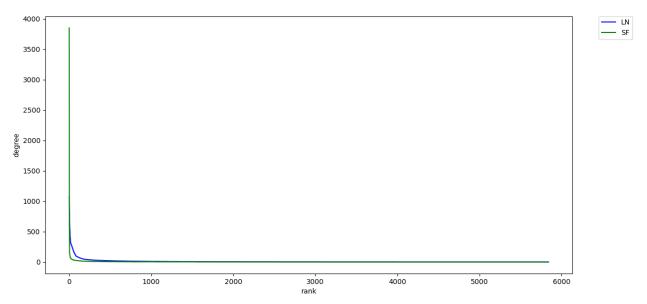
Andando ad analizzare questi due valori possiamo notare che $\lambda \approx 1$ e $\Upsilon > 1$ possiamo quindi già concludere che la LN è una *small world network*.

Ciò viene confermato anche dal sigma (Υ / λ) che è uguale a 164.

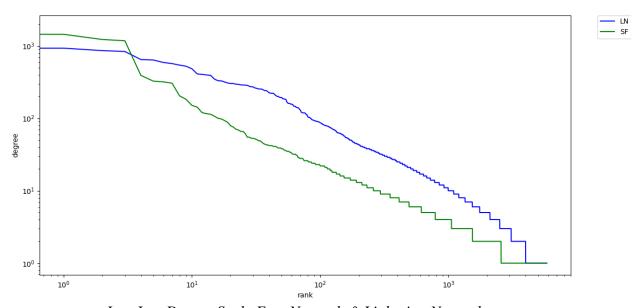
• *Scale-free network*: Una *Scale-free network* è caratterizzata dalla presenza di nodi che fungono da hubs e hanno un grado molto più alto rispetto agli altri.

Quando un nodo deve stabilire un nuovo collegamento, preferisce farlo verso un nodo che ne ha già molti, portando questi ad una crescita esponenziale con l'aumentare del numero dei collegamenti della rete.

In questo caso andremo a confrontare la distribuzione dei *Degree* tra una *Scale-free network* e la LN.



Linear Degree Scale Free Network & Lightning Network



Log-Log Degree Scale Free Network & Lightning Network

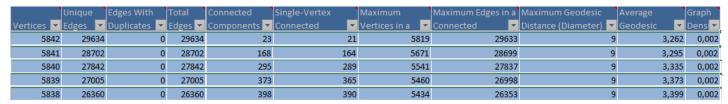
Come si può notare dai grafici precedenti l'andamento della LN si adatta bene a quello di una rete *Scale Free*.

Anche nella LN, infatti, i nodi hanno un incentivo ad aprire canali con nodi altamente connessi, raggiungendo in questo modo una parte maggiore della rete tramite un minor numero di step.

Criticità della rete

La poca densità della rete e la presenza di hubs con posizioni di intermediazione fondamentale per la connessione la rendono vulnerabile a ipotetici attacchi rivolti agli attori che hanno figure centrali.

Di seguito effettueremo un confronto tra la rimozione di nodi centrali con la rimozione di nodi casuali.

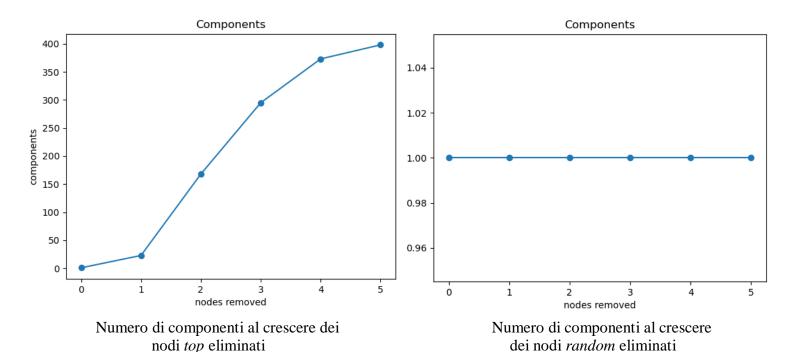


NodeXL - Group Metrics

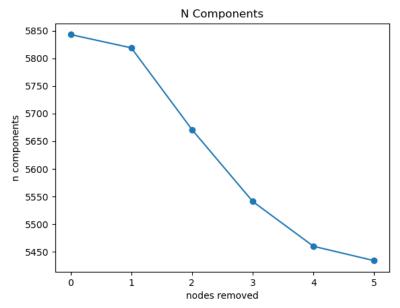
Le statistiche generali restano uguali eccetto per il numero di componenti e come vedremo in seguito per il coefficiente di clustering.

Andando ad aumentare i componenti, infatti, il grafico diventa disconnesso quindi, anche se le altre statistiche restano invariate, la rete viene frammentata non essendo più possibile la connessione tra tutti i nodi.

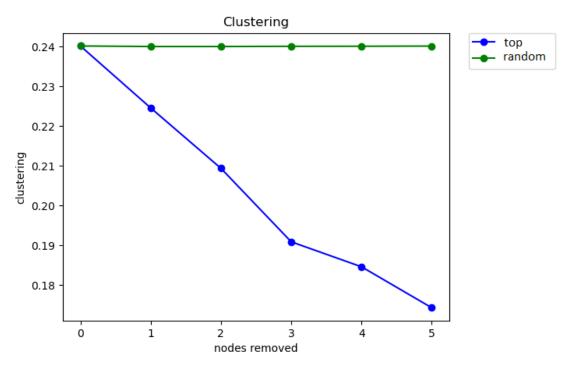
In questo modo singoli nodi o gruppi di nodi più periferici restano isolati dalla rete principale.



17



Numero di nodi nel componente più grande al crescere dei top nodi eliminati



Average Clustering Coefficent al crescere dei nodi eliminati

Andando ad analizzare i precendenti grafici possiamo notare, come ci aspettavamo, che eliminando i nodi più centrali la rete diventa frammentata e il coefficienete medio di clustering diminuisce rendendo, quindi, non tutti i nodi raggiungibili tra loro.

Effettuando, quindi, attacchi mirati su questi tipi di nodi possiamo ostacolare il normale flusso monetario all'interno della LN.

Topology-Based Attacks

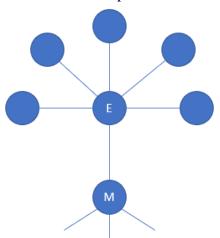
In base alle analisi effettuate precendentemente possiamo, quindi, ipotizzare alcuni tipi di attacco a cui la LN sarebbe vulnerabile:

- **DoS**: un attacco di questo tipo su nodi che fungono da intermediatori potrebbe bloccare il normale flusso di pagamenti all' interno della LN. Ciò potrebbe essere effettuato solo da un nodo ragionevolmente forte (che riesca a re-instradare un grande numero di pagamenti in entrata nel suo nodo verso un nodo bersaglio).
- **Esaurimento del canale**: come detto, ogni canale nella LN ha una certa capacità. Un eventuale nodo, con fondi sufficiente, potrebbe esaurire i fondi di un attore che ha un ruolo centrale nella rete andando anche in questo caso a bloccare il flusso di pagamenti.
- **Isolamento del nodo**: l'attacco di esaurimento verso un nodo centrale potrebbe isolare uno o più nodi collegati ad esso, rendendo il nodo bersaglio incapace di instradare i pagamenti in uscita.

Di seguito illustreremo nello specifico i vari tipi di attacchi:

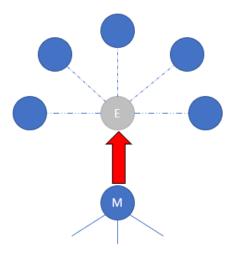
DoS

Supponiamo che il nodo malevolo M abbia una posizione strategica, richieda una commissioni molto bassa o pari a zero e abbia numero di canali aperti sufficiente.



Grazie alla possibilità dei nodi di instradare i pagamenti tramite path che attraversano nodi con commissioni più basse e la posizione strategica di M, quest'ultimo nodo riceverà in entrata un grande flusso di pagamenti.

Il nodo M potrebbe indirizzare a sua volta questo grande numero di pagamenti verso il nodo bersaglio effettuando un attacco DoS.



In questo modo il nodo bersaglio verrebbe estromesso dal flusso di pagamenti comportando un problema sia per il nodo stesso che per la connettività della LN.

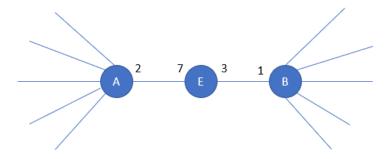
L'attacco può essere ulteriormente amplificato sfruttando il meccanismo di ritardo del *hash time-locked contracts*.

Se l'aggressore smette di partecipare durante le transazioni, allora gli altri nodi potrebbero aver già bloccato le loro criptovalute nel *hash time-locked contracts* e saranno, quindi, in grado di liberarle solo dopo il lock time.

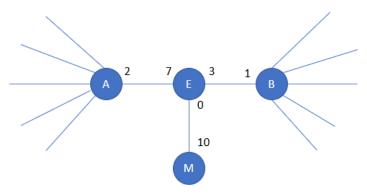
In questo modo il denaro dei nodi del percorso sarà bloccato per un tempo più lungo, impedendo al nodo sorgente di eseguire un'altra transazione tramite un altro percorso.

Esaurimento dei canali & Isolamento nodi

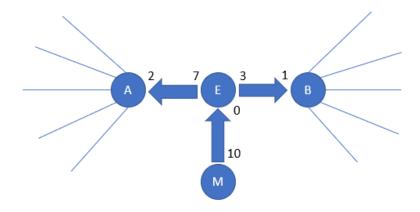
Supponiamo che il nodo E abbia un totale di 10 BTC sulla blockchain. E apra 2 canali con A e B dove deposita rispettivamente 7 e 3 BTC.



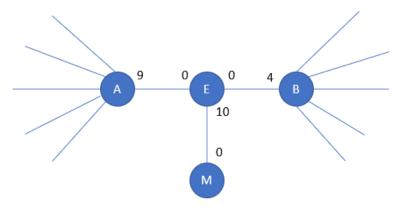
Supponiamo adesso che un nodo malevolo M apra un channel con E depositando nel canale 10 BTC (pari al bilancio totale dei canali in uscita di E).



Il nodo M instraderà nella direzione di B uno o più pagamenti per un totale di 3BTC e verso A per un totale di 7BTC.



In questo modo il nodo E non avrà più fondi per instradare pagamaneti in uscita.



In questa situazione possono verificarsi due situazioni

- sia il nodo E che il nodo M possono decidere di chiudere il canale, che restituirebbe i fondi al nodo E on-chain. In questo modo però E non potrebbe utilizzare i fondi per l'intero periodo che trascorre tra la chiusura del canale e il mining del blocco contenente quest'ultima.
- E chiude il canale unilateralmente. In questo caso, prima che i fondi siano restituiti on-chain, oltre al tempo richiesto nella situazione precedente, si aggiunge l'attesa del lock time. Infatti la chiusura unilaterale comporta la sospensione fino alla scadenza del lock time del hash timelocked contracts.

In entrambi i casi il nodo E, per l'intero periodo d'attesa, non sarà in grado di instradare pagamenti, impendendo quindi i passaggi monetari da A (e tutti i nodi ad esso collegati) a B(e tutti i nodi a esso colegati) e viceversa.

Conclusioni

In conclusione possiamo vedere la LN, da un lato, come un network di trasferimento *off-chain* che va a sopperire ad alcune delle criticità principali della *Blockchain*, dall'altro, come un *network* che data la parziale centralità perde la sicurezza presente sulla catena principale.

Come abbiamo visto, infatti, nella LN ci sono attori centrali dal quale passano la maggior parte dei pagamenti e senza i quali la rete diventa frammentata.

Eseguire un attacco su uno di questi nodi non richiede una potenza computazionale elevata come negli attacchi rivolti alla Blockchain e nemmeno la necessità di avere una quantità molto grande di criptovalute.

Ciò è possibile grazie allo "sfruttamento" sia dagli algoritmi di *routing*, per minimizzare le *fees*, che delle proprietà degli *hash time locked contracts*.

Sitografia

- https://www.docenti.unina.it/webdocenti-be/allegati/materiale-didattico/72957
- https://arxiv.org/pdf/1904.10253.pdf
- https://1ml.com
- https://iol.unibo.it