

Αποκωδικοποίηση δεδομένων και ανίχνευση κοινοτήτων στο δίκτυο Bitcoin.

Δημήτριος Ζέρβας

Διπλωματική Εργασία

Επιβλέπων: Ν. Μαμουλής

Ιωάννινα, Μάρτιος 2022



**ΤΜΗΜΑ ΜΗΧ. Η/Υ & ΠΛΗΡΟΦΟΡΙΚΗΣ
ΠΑΝΕΠΙΣΤΗΜΙΟ ΙΩΑΝΝΙΝΩΝ**

**DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING
UNIVERSITY OF IOANNINA**

Ευχαριστίες

Θα ήθελα να ευχαριστήσω τους επιβλέποντες, τον καθηγητή κ. Μαμουλή και την υποψήφια διδάκτορα κα. Κοσυφάκη για την βοήθεια και την καθοδήγηση τους στην εκπόνηση της διπλωματικής μου εργασίας καθώς και την οικογένεια μου για την στήριξη.

10/03/2022

Δημήτριος Ζέρβας

Περίληψη

Το Bitcoin είναι ένα αποκεντρωμένο ψηφιακό νόμισμα που δημιουργήθηκε και κυκλοφόρησε ως λογισμικό ανοιχτού κώδικα το 2009 από τον Satoshi Nakamoto. Το εν λόγω ψηφιακό νόμισμα, δεν βασίζεται σε κεντρικό διακομιστή για την επεξεργασία και την διεκπεραίωση των συναλλαγών. Αντίθετα, το δίκτυο Peer-to-Peer ρυθμίζει τα Bitcoins, την δημιουργία τους και τις συναλλαγές, σύμφωνα με τη συναίνεση στο λογισμικό δικτύου. Οι συναλλαγές, επαληθεύονται από τους κόμβους δικτύου μέσω της χρήσης κρυπτογραφίας και καταγράφονται σε ένα δημόσιο κατανεμημένο σύστημα, που ονομάζεται Blockchain. Το Bitcoin, εμφανίστηκε ως μια εναλλακτική μορφή χρήματος και έδωσε λύση στην επιβολή περιορισμών του οικονομικού συστήματος (συγκεκριμένοι διαμεσολαβητές ελέγχουν την έκδοση λογαριασμών και τη διεκπεραίωση των συναλλαγών), πραγματοποιώντας ανώνυμες συναλλαγές και παρέχοντας ιδιωτικότητα. Οι χρήστες μπορεί να έχουν μία ή περισσότερες διευθύνσεις για να λαμβάνουν και να στέλνουν ποσά μεταξύ τους (flow), δημιουργώντας ένα δίκτυο, τις κοινότητες (communities). Για την διεκπεραίωση των συναλλαγών μεταξύ των χρηστών, υπάρχουν κάποια πεδία: οι εισροές που προσδιορίζουν την διεύθυνση του αποστολέα και το ποσό που στάλθηκε, οι εκροές που προσδιορίζουν την αντίστοιχη διεύθυνση του παραλήπτη και το ποσό που ελήφθη καθώς και η χρονοσφραγίδα (timestamp) που καθορίζει την χρονική στιγμή που πραγματοποιήθηκε η συναλλαγή.

Στόχος της διπλωματικής εργασίας, είναι η αποκωδικοποίηση των συναλλαγών σε ανώνυμους λογαριασμούς του δικτύου Bitcoin (Bitcoin Core), καθώς και η συλλογή πληροφοριών από τον εντοπισμό και την ανάλυση των κοινοτήτων που συνθέτουν το δίκτυο. Πιο συγκεκριμένα, μετά την εγκατάσταση του δικτύου Bitcoin Core και αποκρυπτογράφοντας τις συναλλαγές μεταξύ των χρηστών, πραγματοποιείται ομαδοποίηση των διευθύνσεων που ανήκουν στον ίδιο χρήστη και εκ νέου ταυτοποίηση του κάθε χρήστη (διαφορετικές διευθύνσεις ανήκουν στον ίδιο χρήστη). Ακόμη, σε κάθε χρήστη ανατίθεται ένα μοναδικό αναγνωριστικό και για κάθε συναλλαγή προστίθεται μια ακμή από τον αποστολέα στον παραλήπτη δημιουργώντας το δίκτυο, και συνεπώς σχηματίζονται κοινότητες με διαφορετικά χαρακτηριστικά. Στην συνέχεια, πραγματοποιείται ανίχνευση κοινοτήτων μέσα στο δίκτυο εφαρμόζοντας γνωστούς αλγορίθμους όπως π.χ. Girvan-Newman κ.α. Η διαδικασία αυτή πραγματοποιείται διότι ενδιαφερόμαστε να ομαδοποιήσουμε τους χρήστες οι οποίοι έχουν κοινά χαρακτηριστικά με σκοπό να εξάγουμε διάφορα συμπεράσματα. Επιπροσθέτως, γίνεται σύγκριση μεταξύ των αλγορίθμων που χρησιμοποιούνται για την ανίχνευση κοινοτήτων ως προς την αποτελεσματικότητά τους. Προκειμένου να αξιολογήσουμε την μέθοδο που υλοποιήσαμε, διεξάχθηκαν πειράματα

χρησιμοποιώντας πραγματικά δεδομένα από το Bitcoin. Τα αποτελέσματα δείχνουν μεγάλη αύξηση στις συναλλαγές που γίνονται μεταξύ των χρηστών με την πάροδο των χρόνων. Επιπροσθέτως, παρατηρείται αύξηση και στον αριθμό των χρηστών και κατά συνέπεια αυξάνεται και ο αριθμός των κοινοτήτων που δημιουργούνται μέσα στο σύστημα. Είναι εμφανές ότι το Bitcoin έχει μπει για τα καλά στην οικονομία ως ένα ευρέως χρησιμοποιούμενο εναλλακτικό ιδιωτικό νόμισμα. Αξίζει να σημειωθεί ότι το 2019, το Ελ Σαλβαδόρ το αναγνώρισε ως επίσημο νόμισμα.

Λέξεις Κλειδιά: Bitcoin, Blockchain, Satoshi Nakamoto, Δίκτυο Peer-to-Peer, Communities, Community Detection, Girvan-Newman, Flow, Timestamp

Abstract

Bitcoin is a decentralized digital currency created and released as open-source software in 2009 by Satoshi Nakamoto. This digital currency is not based on a single administrator for processing transactions. Instead, Peer-to-Peer network regulates Bitcoins, their creation and transactions, according to the network software. These transactions are verified by network nodes using cryptography and recorded in a public distributed system, called Blockchain. Bitcoin used as an alternative form of money and provided a solution to the restrictions of the financial system (certain intermediaries control the issuance of accounts and processing transactions) by conducting anonymous transactions, providing privacy. Users can have one or more addresses to receive and send money to each other (flow) creating a network, called communities. There are several fields for handling transactions between users: the inputs that specify the address of sender and value that was sent, the outputs that specify the corresponding address of recipient and value that was received and timestamp that specifies the time when transaction took place.

The aim of this dissertation is decoding transactions in anonymous accounts of Bitcoin network (Bitcoin Core), as well as to collect information from the identification and analysis of the communities that make up the network. More specifically, after installing the Bitcoin Core network and decrypting the transactions between the users, a grouping is performed between addresses and re-identification of each user is applied (multiple addresses belong to same user). Also, each user is assigned a unique ID and for each transaction an edge is added from sender to recipient creating the network, forming communities with different characteristics. Then, community detection is applied within the network using known algorithms such as, for example, Girvan-Newman etc. This process is made because we are interested in grouping users with common features in order to draw different conclusion. In addition, a comparison is done between the algorithms used to detect communities in terms of their effectiveness.

In order to evaluate the method, we implemented experiments were performed using real data from Bitcoin. The results show a large increase in transactions between users over time. In addition, there is an increase in the number of users and consequently the number of communities created within the system also increases. It is clear that Bitcoin replaced fiat money and used as a widely alternative private currency. It is worth noting that in 2009, El Salvador used it as an official currency.

Keywords: Bitcoin, Blockchain, Satoshi Nakamoto, Peer-to-Peer Network, Communities, Community Detection, Girvan-Newman, Flow, Timestamp

Πίνακας περιεχομένων

Κεφάλαιο 1. Εισαγωγή	1
1.1 Αντικείμενο της εργασίας	1
1.2 Οργάνωση της εργασίας	2
Κεφάλαιο 2. Σχετικές εργασίες	3
Κεφάλαιο 3. Επισκόπηση του Bitcoin	4
3.1 Ανασκόπηση του Blockchain	4
3.1.1 Δίκτυο μεταξύ χρηστών (Peer-to-Peer)	5
3.1.2 Απόδειξη εργασίας (Proof of Work)	6
3.1.3 Αμετάβλητο μητρώο (Immutable ledger)	7
3.2 Τρόπος λειτουργίας του Bitcoin	8
3.2.1 Αρχιτεκτονική δημόσιου κλειδιού	8
3.2.2 Συναλλαγές (Transactions)	9
3.2.2.1 Εκροές (Outputs)	9
3.2.2.2 Εισροές (Inputs)	10
3.2.2.3 Συναλλαγή δημιουργίας (Coinbase transaction)	10
3.2.3 Τύποι σεναρίων (Scripts)	11
3.2.3.1 Pay to Public Key Hash – P2PKH	14
3.2.3.2 Pay to Public Key – P2PK	16
3.2.3.3 Pay to Script Hash – P2SK	17
3.2.3.4 Pay to Witness Public Key Hash – P2WPKH	17
3.2.3.5 Pay to Witness Script Hash – P2WSH	18
3.2.3.6 Τελεστής επιστροφής (OP Return)	18
3.2.4 Εξόρυξη των blocks (Mining)	18
3.2.4.1 Κεφαλαιοποίηση Bitcoin	19
3.2.4.2 Παζλ κατακερματισμού (hash)	20
3.2.4.3 Βαθμός δυσκολίας	20
Κεφάλαιο 4. Σχεδίαση & Υλοποίηση	21
4.1 Εγκατάσταση Bitcoin Core – Full Node	21
4.1.1 Επαλήθευση λήψης του Bitcoin Core	21

4.1.2	Εγκατάσταση Bitcoin Core	22
4.2	Αποκωδικοποίηση δεδομένων	24
4.2.1	Δομή block.....	25
4.2.2	Δεδομένα συναλλαγών.....	29
4.2.3	Άγνωστες διευθύνσεις εισόδου	31
4.2.4	Επαναπροσδιορισμός χρήστη	34
4.3	Χρήση αποκωδικοποιητή	35
4.3.1	Εκτέλεση κώδικα για την αποκωδικοποίηση.....	35
4.3.2	Ομαδοποίηση των διευθύνσεων	38
Κεφάλαιο 5.	Ανίχνευση κοινοτήτων	40
5.1	Αλγόριθμοι ανίχνευσης κοινοτήτων	42
5.1.1	Girvan-Newman	42
5.1.2	Louvain.....	42
5.1.3	Leiden.....	43
5.1.4	Walktrap.....	43
Κεφάλαιο 6.	Πειραματική Αξιολόγηση	44
6.1	Προγραμματιστικά εργαλεία	44
6.2	Ανάλυση αποτελεσμάτων	44
Κεφάλαιο 7.	Επίλογος.....	50

Κεφάλαιο 1. Εισαγωγή

1.1 Αντικείμενο της εργασίας

Το 2008, η παγκόσμια χρηματοπιστωτική κρίση είχε ως αποτέλεσμα την ύφεση και την απογοήτευση απέναντι στο παγκόσμιο χρηματοπιστωτικό σύστημα. Η κρίση αυτή, δημιουργήσε την αμφισβήτηση των ανθρώπων για την προστασία του πλούτου τους αλλά και της πραγματοποίησης επενδύσεων χωρίς τον έλεγχο και περιορισμό των κυβερνήσεων. Κατά συνέπεια, δημιουργήθηκε η ανάγκη για εναλλακτικές μορφές χρήματος, ενός ιδιωτικού νομίσματος, το οποίο δεν θα βασίζεται σε κάποιο ενδιάμεσο διακομιστή, όπως κράτος ή κεντρική τράπεζα, για την ολοκλήρωση μιας συναλλαγής. Το κρυπτονόμισμα Bitcoin, αποτελεί το πρώτο ψηφιακό ιδιωτικό νόμισμα. Από τον Ιανουάριο του 2018, είναι το πιο ευρέως χρησιμοποιούμενο εναλλακτικό νόμισμα (από το παραστατικό χρήμα), με συνολικό ανώτατο όριο αγοράς, περίπου 836 δισεκατομμύρια δολάρια.

Το Bitcoin [Naka08] κυκλοφόρησε το 2009 από τον Satoshi Nakamoto και έδωσε λύση στην ύπαρξη περιορισμών του οικονομικού συστήματος, καθώς δεν βασίζεται σε κάποια ενδιάμεση αρχή. Το δίκτυο Peer-to-Peer [1] είναι υπεύθυνο για την διεκπεραίωση των συναλλαγών και την παραγωγή των Bitcoins, σύμφωνα με την συναίνεση στο λογισμικό δικτύου. Οι κόμβοι (δηλαδή οι χρήστες) που είναι συνδεδεμένοι στο δίκτυο Bitcoin, είναι υπεύθυνοι για την επαλήθευση, την αξιοπιστία και την επικύρωση των συναλλαγών μέσω του μηχανισμού Proof of Work [2]. Με αυτόν τον τρόπο, το Bitcoin είναι πλήρως αποκεντρωμένο. Ακόμη, βασίζεται στις αρχές της ισχυρής κρυπτογραφίας, δίνοντας την δυνατότητα στους χρήστες να πραγματοποιούν ασφαλείς ανώνυμες συναλλαγές. Ωστόσο, παρόλο που το Bitcoin παρέχει ανωνυμία, οι συναλλαγές του Bitcoin αποθηκεύονται στη δημόσια τεχνολογία Blockchain [3] και είναι δυνατή η ιχνηλάτηση του ιστορικού των συναλλαγών.

Στο δίκτυο Bitcoin συνυπάρχουν χρήστες ποικίλων χαρακτηριστικών που συμβάλουν στον σχηματισμό ομάδων και πιο συγκεκριμένα κοινοτήτων. Οι κοινότητες αυτές, αποτε-

λούνται από άτομα που μοιράζονται ορισμένα κοινά χαρακτηριστικά και ιδιότητες με αποτέλεσμα την συχνότερη αλληλοεπίδραση ανάμεσα τους συγκριτικά με το υπολειπόμενο δίκτυο. Στόχος της διπλωματικής εργασίας είναι η αποκωδικοποίηση του δικτύου Bitcoin καθώς και η ανίχνευση κοινοτήτων μέσα στο δίκτυο. Πιο συγκεκριμένα, ο εντοπισμός κοινοτήτων είναι πολύ σημαντικός διότι βοηθά στην εξαγωγή διάφορων συμπερασμάτων λαμβάνοντας υπόψιν τα χαρακτηριστικά που απαρτίζουν την κάθε κοινότητα. Για παράδειγμα, στο δίκτυο Bitcoin υπάρχουν περιπτώσεις όπου σχηματίζεται μια κοινότητα, στην οποία οι χρήστες ανταλλάσσουν μεγάλα ποσά σε σχέση με το υπόλοιπο δίκτυο και ίσως αυτό παραπέμπει σε ύποπτες συναλλαγές, όπως αγορά παράνομων αγαθών και υπηρεσιών ή ξέπλυμα χρήματος. Χαρακτηριστικό παράδειγμα είναι η περίπτωση της διαδικτυακής μαύρης αγοράς Silk Road [4], που χρησιμοποιήθηκε για την πώληση παράνομων προϊόντων. Το Ομοσπονδιακό Γραφείο Ερευνών της Αμερικής (Federal Bureau of Investigation – FBI), ανέλυσε το δίκτυο συναλλαγών του Bitcoin μέσω διάφορων κοινοτήτων και συγκεκριμένων μοτίβων, όπου μέσω της ιχνηλάτησης των συναλλαγών εξακριβώθηκε και διακόπηκε η σύνδεση στην ιστοσελίδα Silk Road.

1.2 Οργάνωση της εργασίας

Η συγκεκριμένη διπλωματική εργασία αποτελείται από 7 κεφάλαια. Πιο αναλυτικά, στο κεφάλαιο 2 γίνεται περιγραφή εργασιών που είναι σχετικές με την ομαδοποίηση των διευθύνσεων που ανήκουν στον ίδιο χρήστη και για την διαδικασία της ανίχνευσης κοινοτήτων. Στο κεφάλαιο 3 παρουσιάζεται η αρχιτεκτονική και ο τρόπος λειτουργίας του Bitcoin καθώς και η διαδικασία της εξόρυξης που πραγματοποιείται ο τρόπος επικύρωσης και διεκπεραίωσης των συναλλαγών μέσω του μηχανισμού συναίνεσης Proof of Work. Στο κεφάλαιο 4 αναλύονται οι λεπτομέρειες για την σχεδίαση και υλοποίηση της ιχνηλάτησης και αποκωδικοποίησης των συναλλαγών. Πιο συγκεκριμένα, αναλύεται η διαδικασία εγκατάστασης του δικτύου Bitcoin, η αποκωδικοποίηση των δεδομένων καθώς και ο επαναπροσδιορισμός των διευθύνσεων του κάθε χρήστη. Στην συνέχεια, γίνεται αναφορά στον τρόπο χρήσης του αποκωδικοποιητή και στην εκτέλεση του κώδικα. Στο κεφάλαιο 5 περιγράφεται η διαδικασία ανίχνευσης κοινοτήτων όπως επίσης και οι αλγόριθμοι που χρησιμοποιούνται για την ομαδοποίηση των χρηστών με παρόμοια χαρακτηριστικά. Στο κεφάλαιο 6 παρουσιάζονται τα πειράματα που πραγματοποιήθηκαν με τα αποτελέσματα τους για την ανάλυση του δικτύου και την ανίχνευση κοινοτήτων. Κλείνοντας, στο κεφάλαιο 7 αναφέρονται πιθανές επεκτάσεις καθώς και τα συμπεράσματα των πειραμάτων.

Κεφάλαιο 2. Σχετικές εργασίες

Το Bitcoin σχεδιάστηκε και υλοποιήθηκε ως μια εναλλακτική μορφή χρήματος και έδωσε λύση στους περιορισμούς οικονομίας της αγοράς, παρέχοντας ανωνυμία στις συναλλαγές. Παρόλα αυτά, οι συναλλαγές του Bitcoin αποθηκεύονται στην δημόσια τεχνολογία Blockchain και είναι δυνατή η ιχνηλάτηση του ιστορικού των συναλλαγών σε μια χρονική στιγμή, των διευθύνσεων μαζί με το συνολικό ποσό που ανταλλάσσουν. Ως αποτέλεσμα, ορισμένοι χρήστες κατέχουν παραπάνω από μία διεύθυνση, περιορίζοντας την διαδικασία της ιχνηλάτησης, διατηρώντας την ιδιωτικότητα.

Οι Remy κ.α. δημοσίευσαν ένα άρθρο [ReRM10] περιγράφοντας κάποια υδριβικά για τον επαναπροσδιορισμό πολλαπλών διευθύνσεων που ανήκουν στον ίδιο χρήστη. Σύμφωνα με την αρχιτεκτονική του δικτύου Bitcoin (Βλ. 3.2.1), το σύνολο των διευθύνσεων που περιέχονται στις εισροές κάθε συναλλαγής αντιστοιχούν στον ίδιο χρήστη, όπως επίσης και η αλλαγή διεύθυνσης που περιέχεται στις εκροές και αναγνωρίζεται με ένα συγκεκριμένο τρόπο. Στην συνέχεια, στο δίκτυο αυτό εφαρμόζεται ένας αλγόριθμος ανίχνευσης κοινοτήτων, καθώς οι κοινότητες αντιστοιχούν σε μοναδικούς χρήστες.

Οι Mao κ.α. [MaZh17] κατηγοριοποιούν τον κάθε χρήστη στο δίκτυο χρησιμοποιώντας αλγόριθμους κατηγοριοποίησης όπως K-Means, Node2Vector και Fiedler Vector Method. Κατά την ταξινόμηση των χρηστών σε διάφορες συστάδες (clusters), προκύπτουν κάποια συμπεράσματα για την δομή κοινοτήτων και την ανάλυση ροής των νομισμάτων, καθώς κάθε συστάδα περιέχει χρήστες με κοινά χαρακτηριστικά. Για παράδειγμα, η συστάδα 1 ομαδοποιεί τους χρήστες που ασχολούνται με την εξόρυξη και ανταλλάσσουν τα Bitcoins αυτά με μετρητά, η συστάδα 2 αποτελείται από εμπόρους (αποδέχονται τα Bitcoins ως νόμισμα, όπως ιστοσελίδες τζόγου ή ιστοσελίδες για δωρεές) κ.α.

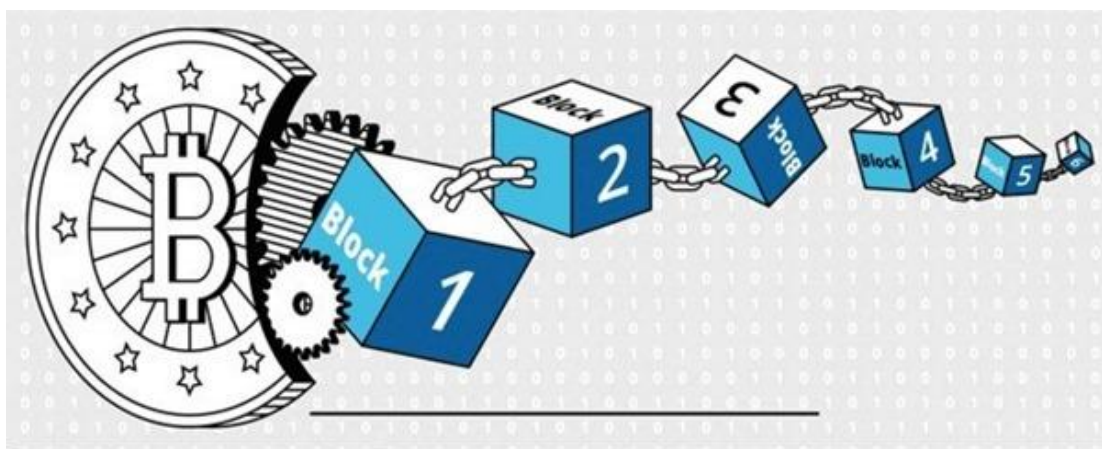
Οι Yu κ.α. [YuBu15] περιγράφουν το απόρρητο του πρωτοκόλλου Bitcoin και των συναλλαγών καθώς και την διαδικασία ανίχνευσης και ανάλυσης των κοινοτήτων που συνθέτουν το δίκτυο. Μετά την δημιουργία του δικτύου και την εκ νέου ταυτοποίηση του κάθε χρήστη, με την χρήση εργαλείων ανάλυσης δικτύου, αναλύονται κοινότητες για την εύρεση χρηστών που αλληλοεπιδρούν συχνά μεταξύ τους. Με αυτό τον τρόπο αντιστοιχίζονται ταυτότητες σε διευθύνσεις Bitcoin και επίσης προσδιορίζονται μοτίβα χρήσης του Bitcoin, εφαρμόζοντας τον αλγόριθμο του Girvan-Newman, και του Louvain.

Κεφάλαιο 3. Επισκόπηση του Bitcoin

Στο συγκεκριμένο κεφάλαιο αναλύεται η αρχιτεκτονική, ο τρόπος παραγωγής και λειτουργίας του Bitcoin. Περιγράφεται η διαδικασία εξόρυξης των blocks που περιέχουν τις συναλλαγές, η επικύρωση και η διεκπεραίωση των συναλλαγών μέσω του μηχανισμού συναίνεσης Proof of Work, η αναμετάδοση τους στο δίκτυο μέσω του συστήματος Peer-to-Peer καθώς και τα βασικά χαρακτηριστικά της τεχνολογίας Blockchain.

3.1 Ανασκόπηση του Blockchain

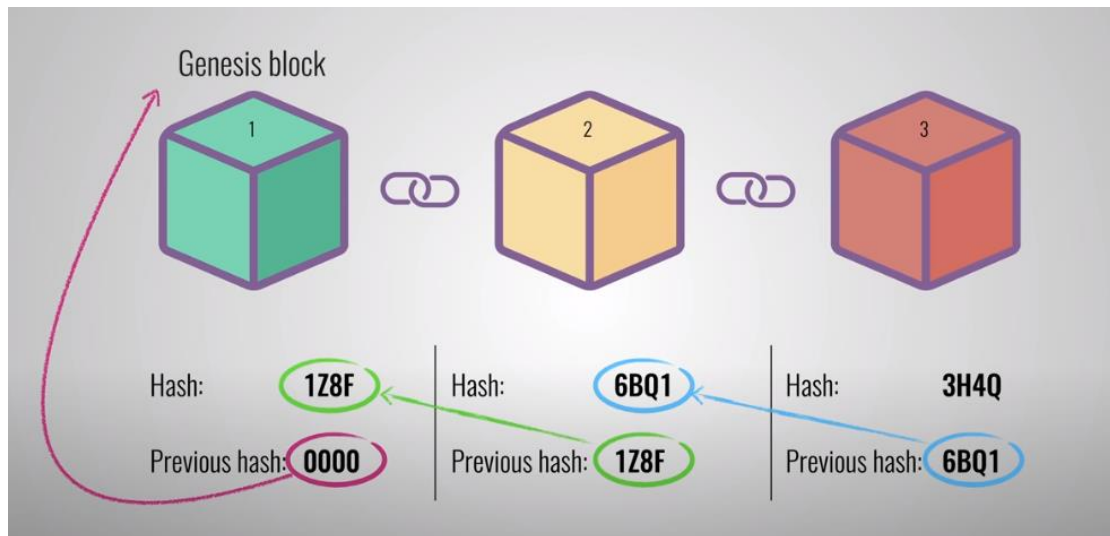
Η τεχνολογία Blockchain (ή αλυσίδα μπλοκ ή μπλοκ αλυσίδας), είναι ένα μητρώο (ledger) ή μια βάση δεδομένων, στην οποία αποθηκεύονται και επαληθεύονται δεδομένα βασισμένα στην κρυπτογραφία. Οι πληροφορίες αυτές συνδέονται άμεσα μεταξύ τους, σχηματίζοντας μία συνεχή αλυσίδα δεδομένων, τα μπλοκ (blocks).



Εικόνα 1: Blockchain (ή μπλοκ αλυσίδας) [5].

Κάθε block έχει 3 βασικά χαρακτηριστικά:

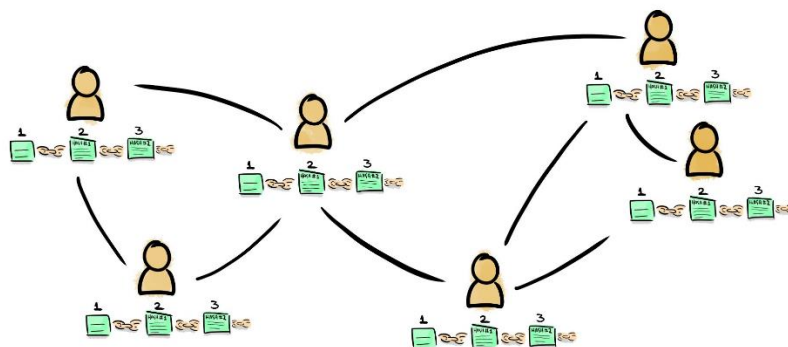
1. την πληροφορία που καταγράφεται (είτε μία συναλλαγή, είτε ένα συμφωνητικό)
2. ένα μοναδικό αλφαριθμητικό που χρησιμοποιείται σαν αναγνωριστικό του κάθε block (hash ή hash id ή block hash)
3. το hash του αμέσως προηγούμενου block (previous hash ή previous hash id ή previous block hash)



Εικόνα 2: Κάθε επόμενο block δείχνει στο hash του αμέσως προηγούμενου block [6]

3.1.1 Δίκτυο μεταξύ χρηστών (Peer-to-Peer)

Η τεχνολογία Blockchain λειτουργεί ως ένα αποκεντρωμένο (decentralized) και διανεμημένο (DLT - distributed ledger technology) μητρώο. Τα δεδομένα συγχρονίζονται ομοτίμα, επαληθεύονται και επικυρώνονται ταυτόχρονα μέσω των συνδεδεμένων κόμβων του δικτύου, οι οποίοι αποθηκεύουν ένα αντίγραφο των δεδομένων. Κάθε κόμβος αντιστοιχεί σε ένα απλό χρήστη, ο οποίος έχει εγκαταστήσει το απαιτούμενο λογισμικό, και μεταδίδει ενημερώσεις με τους άλλους χρήστες. Η ιδιωτικότητα και η προστασία των περιεχομένων του μητρώου είναι γεγονός καθώς δεν απαιτείται η ύπαρξη ενός κεντρικού διακομιστή.



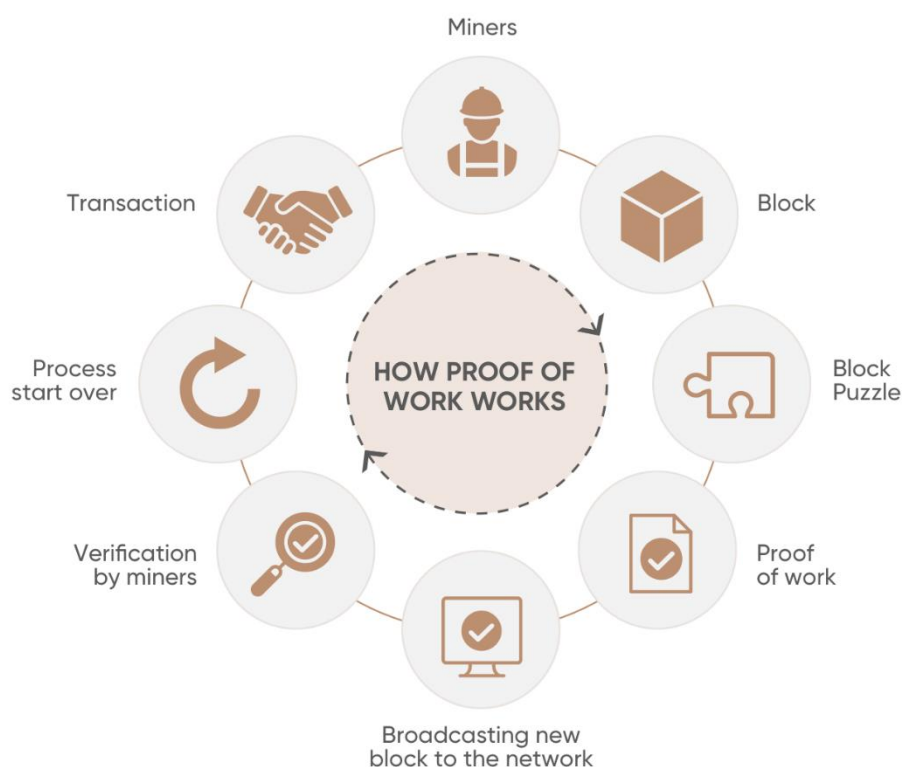
Εικόνα 3: Απεικόνιση χρηστών Bitcoin [7].

Αξιοσημείωτο είναι ότι κάθε block αποθηκεύει δεδομένα σε μια συγκεκριμένη επιτρεπόμενη χωρητικότητα και όταν αυτό είναι πλήρες, τότε δημιουργείται ένα άλλο block που

το ακολουθεί, σχηματίζοντας μια συνεχή αλυσίδα δεδομένων. Προκειμένου να ολοκληρωθεί ένα block συναλλαγών και να ενταχθεί στο Blockchain πρέπει να λυθεί ένας κρυπτογραφημένος γρίφος μιας μαθηματικής πράξης που λέγεται Proof of Work (PoW).

3.1.2 Απόδειξη εργασίας (Proof of Work)

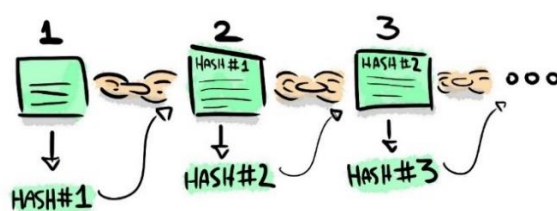
Το Proof of Work είναι πρωτόκολλο ή μηχανισμός που επιτρέπει την συναίνεση σε ένα δίκτυο Blockchain. Πιο συγκεκριμένα, ο μηχανισμός αυτός, συμβάλει στην επίλυση ενός μαθηματικού προβλήματος, τον κρυπτογραφημένο γρίφο (ή παζλ), που προσπαθούν να λύσουν οι κόμβοι του δικτύου (miners) μέσω υπολογιστική ισχύος (hashing). Η διαδικασία είναι γνωστή ως εξόρυξη (mining) και χρησιμοποιείται για την δημιουργία ενός block. Τα παζλ είναι δύσκολα ως προς την επίλυση, αλλά εύκολα ως προς την επαλήθευσή τους. Ο πρώτος miner που θα λύσει το παζλ, θα μεταδώσει το block σε όλο το δίκτυο. Στην συνέχεια, όταν οι υπόλοιποι miners επαληθεύσουν ότι η λύση του είναι σωστή, θα επικυρώσουν το block και ο miner θα λάβει μία ανταμοιβή (reward).



Εικόνα 4: Τρόπος λειτουργίας του Proof of Work [8].

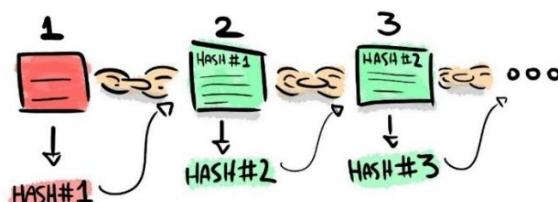
3.1.3 Αμετάβλητο μητρώο (Immutable ledger)

Μια χαρακτηριστική πληροφορία δίπλα στο hash είναι η χρονοσφραγίδα (timestamp), η οποία καθορίζει την χρονική στιγμή που δημιουργήθηκε ένα συγκεκριμένο block. Ως αποτέλεσμα, ο κατακερματισμός και η χρονοσφραγίδα διασφαλίζουν ότι τα blocks είναι άμεσα συνδεδεμένα μεταξύ τους και είναι αδύνατο να παραποιηθεί η αλυσίδα block. Δεδομένου ότι κάθε block περιέχει το hash του προηγούμενου block, εάν τροποποιηθεί κάτι σε ένα block η αλλαγή αυτή θα διαδοθεί σε κάθε block του Blockchain, ακυρώνοντας το block. Για παράδειγμα στην ακόλουθη αλυσίδα όπου έχει υπολογιστεί το hash κάθε block,



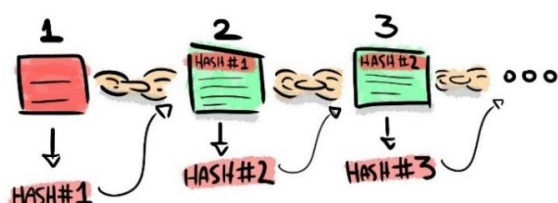
Εικόνα 5: Παράδειγμα μιας αλυσίδας μπλοκ [7].

εάν ένας χρήστης πραγματοποιήσει επίθεση θέλοντας να τροποποιήσει κάποιες πληροφορίες στο block_1, το HASH#1 κατά συνέπεια θα υπολογιστεί ξανά και θα έχει διαφορετικό περιεχόμενο.



Εικόνα 6: Ο επιτιθέμενος χρήστης αλλάζει τις πληροφορίες του HASH#1 [7].

Καθώς το HASH#1 περιέχεται στο block_2, το HASH#2 θα τροποποιηθεί επίσης και η αλλαγή αυτή θα περάσει στο Blockchain, όπου οι κόμβοι του δικτύου θα ακυρώσουν την τροποποίηση του block.



Εικόνα 7: Το Blockchain, δεν μπορεί να τροποποιηθεί ως προς το ιστορικό του [7].

3.2 Τρόπος λειτουργίας του Bitcoin

Το Bitcoin μπορεί να χρησιμοποιηθεί όπως το παραστατικό χρήμα (πχ ευρώ), για οικονομικές συναλλαγές αλλά και για την αγορά φυσικών αγαθών και υπηρεσιών. Η συναλλαγή είναι μια μεταφορά αξίας σε Bitcoin μεταξύ χρηστών, που μεταδίδεται στο δίκτυο και αποθηκεύεται στα blocks. Σύμφωνα με τον Satoshi Nakamoto, το Bitcoin ορίζεται ως ένα ηλεκτρονικό νόμισμα αποτελούμενο από ψηφιακές υπογραφές. Για την ολοκλήρωση μιας συναλλαγής, ο αποστολέας υπογράφει ένα συνδυασμό μεταξύ της εξόδου κατακερματισμού της προηγούμενης συναλλαγής και του δημόσιου κλειδιού του παραλήπτη. Στην συνέχεια, χρησιμοποιείται η υπογραφή του παραλήπτη για την επαλήθευση της ιδιοκτησίας του δημόσιου κλειδιού του και για την πραγματοποίηση της συναλλαγής.

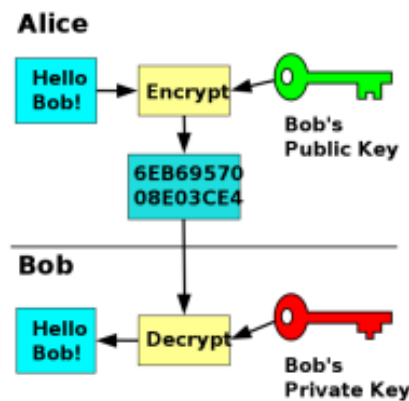
3.2.1 Αρχιτεκτονική δημόσιου κλειδιού

Για την αποθήκευση Bitcoin χρησιμοποιείται ένα ψηφιακό πορτοφόλι (wallet), το οποίο περιέχει ψηφιακές υπογραφές (ή ιδιωτικά κλειδιά). Ωστόσο κανείς δεν κατέχει Bitcoin, παρά μόνο ένα συνδυασμό δημόσιου (public) και ιδιωτικού κλειδιού (private key). Το public και το private key είναι απαραίτητα για την μεταφορά Bitcoin. Το public key αντιστοιχεί σε μία μοναδική διεύθυνση (address) και είναι ορατό σε όλο το δίκτυο, ενώ αντίστοιχα το private key αποδεικνύει την ιδιοκτησία των Bitcoins (μέσω του wallet) και δεν είναι ορατό στο δίκτυο. Το private key είναι ένας τυχαίος αριθμός μεταξύ 2^1 και 2^{256} και χρησιμοποιείται ως υπογραφή ιδιοκτησίας (signature) που απαιτείται για την δαπάνη των Bitcoins. Το public key παράγεται από την εφαρμογή του αλγόριθμου ECDSA [9] στο private key. Πιο συγκεκριμένα, χρησιμοποιούνται συγκεκριμένες διαδικασίες κατακερματισμού και κρυπτογραφίας στο private key με αποτέλεσμα να παραχθεί το public key. Ωστόσο το αντίθετο δεν μπορεί να συμβεί καθώς χρησιμοποιείται ασύμμετρη κρυπτογραφία.



Εικόνα 8: Το ιδιωτικό κλειδί παράγει το δημόσιο κλειδί, αλλά το αντίστροφο δεν ισχύει [10].

Έστω ότι έχουμε τρεις χρήστες τον Bob, την Alice και τον Charlie συνδεδεμένους σε ένα ένα κοινό δίκτυο. Το private key της Alice είναι η ψηφιακή της υπογραφή, αποδεικνύει δηλαδή ότι πραγματοποίησε μια συναλλαγή. Αν η Alice θέλει να στείλει στον Bob ένα μήνυμα μέσω του κοινού δικτύου, χωρίς να έχει πρόσβαση ο Charlie, πρέπει να κρυπτογραφήσει το μήνυμα με το δικό της private key και το public key του Bob και να το στείλει στον Bob. Η Alice δημιουργεί την έξοδο κατακερματισμού (hash output), το μήνυμα της δηλαδή που στάλθηκε στο public key του Bob. Χρησιμοποιώντας το hash output και το private key, ο Bob μπορεί να αποκρυπτογραφήσει και να λάβει το μήνυμα. Ο Charlie ωστόσο δεν μπορεί να διαβάσει το μήνυμα, καθώς έχει το public key της Alice και όχι την υπογραφή της. Πιο απλά, η Alice έχει έναν χαρτοφύλακα, τον οποίο μπορεί να ανοίξει ο Bob που γνωρίζει τον συνδυασμό.



Εικόνα 9: Η Alice υπογράφει με το private key, και ο Bob επαληθεύει ότι έλαβε το μήνυμα της Alice [11].

3.2.2 Συναλλαγές (Transactions)

Οι συναλλαγές αποτελούν το βασικό χαρακτηριστικό λειτουργίας του δικτύου Bitcoin, καθώς όλοι οι μηχανισμοί λειτουργούν για την διασφάλιση της δημιουργίας και εγκυρότητας της συναλλαγής, την διάδοση της στο δίκτυο και την αποθήκευση της στην τεχνολογία Blockchain. Μια συναλλαγή περιέχει υπογεγραμμένες και κρυπτογραφημένες εισροές (τι στάλθηκε) και εκροές (τι ελήφθη) που περιέχουν οδηγίες για την διεκπεραίωση της.

3.2.2.1 Εκροές (Outputs)

Κάθε καινούργια συναλλαγή δημιουργεί εκροές που αποθηκεύονται στο μητρώο. Κάθε φορά που ένας χρήστης λαμβάνει Bitcoin, ουσιαστικά λαμβάνει εκροές. Αυτές οι εκροές,

λέγονται εκροές συναλλαγής που δεν δαπανήθηκαν ακόμη (Unspent Transaction Outputs – UTXO), αναγνωρίζονται απ' όλο το δίκτυο και είναι διαθέσιμες για σπατάλη σε μελλοντικές συναλλαγές. Κάθε UTXO, αποκτά αδιαίρετη τιμή και προσδιορίζεται από το μοναδικό αναγνωριστικό της κάθε συναλλαγής (transaction id).

Οι UTXO αποτελούνται από δύο πεδία:

- > Την αξία σε Satoshi, όπου 100.000.000 Satoshi = 1 BTC (Bitcoin).
- > Το σενάριο κλειδώματος (locking script ή scriptPubKey), που καθιστά το ποσό κλειδωμένο μέχρι να βρεθεί η εισροή που πληρεί τις προϋποθέσεις για το ξεκλείδωμα και την δαπάνη του ποσού αυτού.

3.2.2.2 Εισροές (Inputs)

Οι UTXO που καταναλώνονται λέγονται inputs. Τα inputs, λόγω της αδιαίρετης τιμής, πρέπει να δαπανηθούν εξ' ολοκλήρου. Για παράδειγμα, εάν η διεύθυνση της Alice λαμβάνει 50 BTC και θέλει να στείλει 35 BTC στον Bob, πρέπει να δαπανηθούν 50 BTC. Συνεπώς, ο Bob λαμβάνει 35 BTC και το υπόλοιπο ποσό (change) επιστρέφει στην Alice ως μια νέα συναλλαγή.

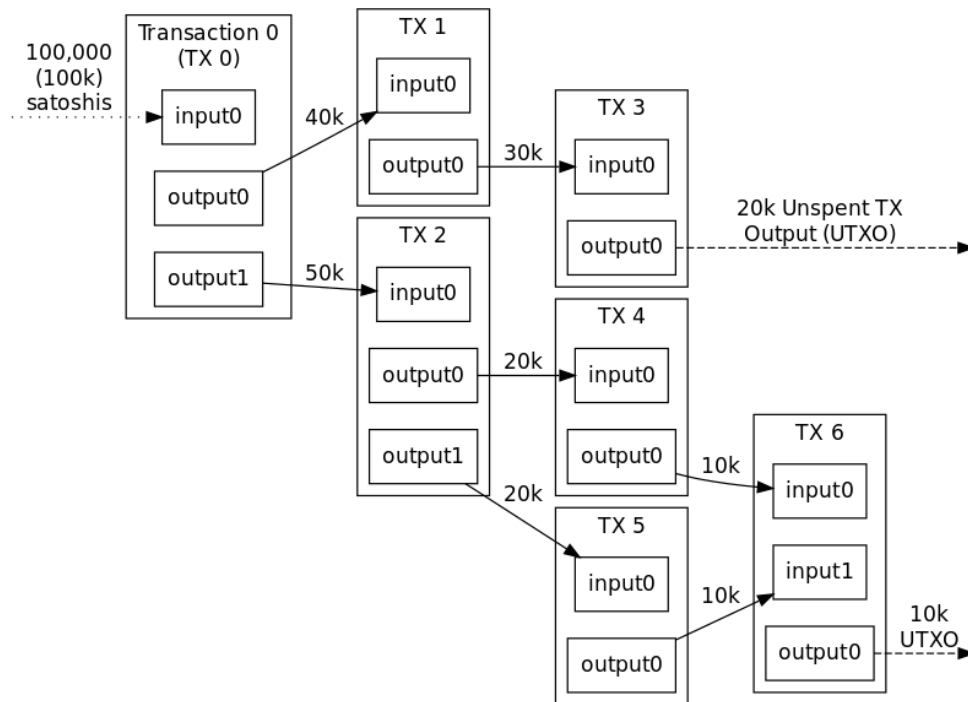
Τα inputs αποτελούνται από τρία πεδία:

- > Το μοναδικό αναγνωριστικό της προηγούμενης συναλλαγής (previous tx id).
- > Τον δείκτη που αναφέρεται σε μία συγκεκριμένη UTXO (previous tx index).
- > Το σενάριο ξεκλειδώματος (unlocking script ή scriptSig), το οποίο ικανοποιεί τις προϋποθέσεις δαπάνης που έχουν οριστεί από την UTXO. Είναι μια υπογραφή που αποδεικνύει την ιδιοκτησία της διεύθυνσης Bitcoin.

Στην περίπτωση που υπάρχουν πολλαπλά inputs σε μια συναλλαγή, το σύνολο αυτών χρησιμοποιείται για την σπατάλη στα outputs.

3.2.2.3 Συναλλαγή δημιουργίας (Coinbase transaction)

Η συναλλαγή δημιουργίας αναφέρεται στην πρώτη συναλλαγή κάθε καινούργιου block που δημιουργείται. Είναι η επιβράβευση του miner για την δημιουργία του block. Αυτή η συναλλαγή έχει μόνο outputs και κανένα input. Κατά αυτόν τον τρόπο, πραγματοποιείται η πίστωση ανταμοιβής στους miners.



Triple-Entry Bookkeeping (Transaction-To-Transaction Payments) As Used By Bitcoin

Εικόνα 10: Παράδειγμα μεταφοράς των UTXO [12].

3.2.3 Τύποι σεναρίων (Scripts)

Για την επαλήθευση μιας συναλλαγής, οι miners που είναι συνδεδεμένοι στο δίκτυο Bitcoin εκτελούν κάποιες εντολές σεναρίων χρησιμοποιώντας την γλώσσα προγραμματισμού Forth [13].

Τα scripts [14] αυτά είναι:

- > το σενάριο κλειδώματος (scriptPubKey ή locking script).
- > το σενάριο ξεκλειδώματος (scriptSig ή unlocking script).

Τα scripts βασίζονται στην λειτουργία στοίβας (stack-based) σε συνδυασμό με το Πολωνικό σύστημα αντιστροφής (Reverse polish notation - RPN) [15], και κάθε εντολή εκτελείται μια φορά ακριβώς.

Παράδειγμα 1. RPN βασισμένο σε στοίβα

Η πρόσθεση 3+4 στο RPN ισοδυναμεί με 34+ και αντίστοιχα οι εντολές σεναρίου είναι οι εξής: OP_3 OP_4 OP_ADD, όπου ο τελεστής OP_ADD πρόσθεσης*.

*Σημείωση: Το πρόθεμα OP κάθε τελεστή είναι μια υπογραφή της γλώσσας script, και προσδιορίζει την εντολή εκτέλεσης του κάθε script.

Η παρακάτω εικόνα περιέχει τα βήματα που απαιτούνται:

- > 1. Ώθηση του 3 στην στοίβα.
- > 2. Ώθηση του 4 στην στοίβα.
- > 3. Λόγω του τελεστή πρόσθεσης, πρώτα θα βγει το 4 και μετά το 3 (LIFO).
- > 4. Το 3 και το 4 θα προστεθούν και το αποτέλεσμα ωθείται στην στοίβα.

Step 1: Pushing 3 into the stack.

3

Step 2: Pushing 4 into the stack

4
3

Step 3: Now we got the addition operation, which will pop 4 and 3 out (in that order, remember LIFO)

So, 4 gets popped first.

3

And then 3.

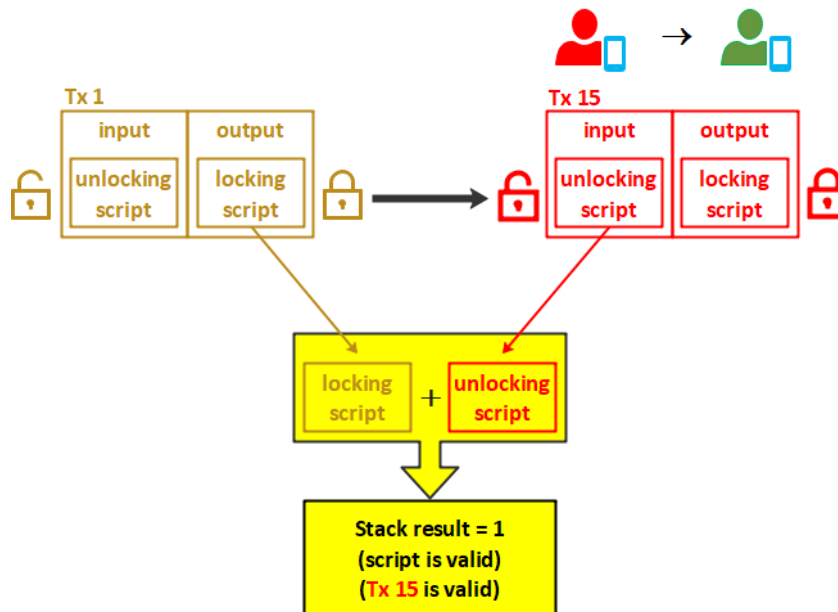
Step 4: 3 and 4 gets added and the result, 7, gets pushed onto the stack.

7

Εικόνα 11: Παράδειγμα RPN [13].

Οι UTXO κλειδώνονται από το locking script ενώ τα inputs περιέχουν το unlocking script. Πιο συγκεκριμένα το locking script δημιουργεί ένα κρυπτογραφημένο παζλ που λύνεται μόνο με το μοναδικό κλειδί ξεκλειδώματος. Με αυτόν τον τρόπο, οι UTXO μπορούν να σπαταληθούν μόνο από τον ιδιοκτήτη της διεύθυνσης Bitcoin. Στην αρχή, εκτελείται το unlocking script. Όταν ολοκληρωθεί χωρίς σφάλματα τότε αποθηκεύεται η τιμή της στοιίβας και εκτελείται το locking script. Μετά την ολοκλήρωση της διαδικασίας, συγκρίνεται με το αποτέλεσμα της στοιίβας και αν η τιμή επιστροφής είναι 1 ή Αληθής (True), τότε το input είναι μια έγκυρη UTXO. Στην παρακάτω εικόνα, για την δαπάνη των inputs της συναλλαγής Tx15, πρέπει να επαληθευτεί ότι ο χρήστης_1 (κόκκινο ανθρωπάκι) έχει λά-

βει UTXO από την Tx1 και είναι έγκυρες. Αν το unlocking script ξεκλειδώσει την UTXO και η τιμή επιστροφής είναι 1, ο χρήστης_1 μπορεί να ξοδέψει την UTXO ως input στην συναλλαγή Tx15.



Εικόνα 12: Αποτέλεσμα στοίβας 1 - Έγκυρη UTXO [14].

Οι τύποι των σεναρίων* κατατάσσονται σε:

- > Πληρωμή σε κατακερματισμένο δημόσιο κλειδί (Pay to Public Key Hash)
- > Πληρωμή σε δημόσιο κλειδί (Pay to Public Key)
- > Πληρωμή σε κατακερματισμένο σενάριο (Pay to Script Hash)
- > Πληρωμή σε κατακερματισμένο δημόσιο κλειδί χωρίς ψηφιακή υπογραφή (Pay to Witness Public Key Hash)
- > Πληρωμή σε κατακερματισμένο σενάριο χωρίς ψηφιακή υπογραφή (Pay to Witness Script Hash)
- > Τελεστής επιστροφής (OP Return)

*Υπενθύμιση: σενάριο κλειδώματος = scriptPubKey ή locking script
 σενάριο ξεκλειδώματος = scriptSig ή unlocking script

3.2.3.1 Pay to Public Key Hash – P2PKH

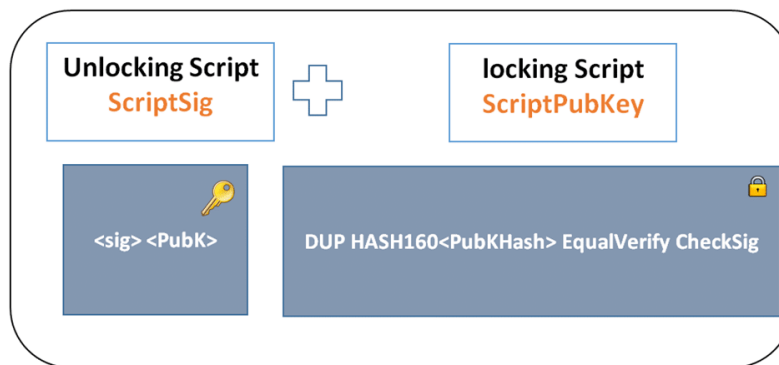
Για την συναλλαγή μιας P2PKH, οι εντολές σεναρίου είναι:

```
> scriptPubKey = OP_DUP OP_HASH160 <public address>  
    OP_EQUALVERIFY OP_CHECKSIG  
  
> scriptSig = <signature> <public key>
```

Παράδειγμα 2. Επιτυχημένη συναλλαγή P2PKH

Όταν η Alice στέλνει μια UTXO στον Bob, ο Bob λαμβάνει και ξεκλειδώνει το μήνυμα χρησιμοποιώντας την υπογραφή (signature) του scriptPubKey. Δηλαδή, σε υλοποίηση στοίβας:

```
> scriptPubKey = OP_DUP OP_HASH160 <Bob's public address>  
    OP_EQUALVERIFY OP_CHECKSIG  
  
> scriptSig = <Bob's signature> <Bob's public key>
```



Εικόνα 13: Locking & unlocking script μιας P2PKH [15].

Για την επαλήθευση της συναλλαγής, τα βήματα είναι:

1. Ώθηση της signature και public key του Bob στην στοίβα.
2. Ο τελεστής εντολής OP_DUP:
 - > Αφαιρεί από την στοίβα το public key του Bob.
 - > Το αντιγράφει και ωθεί το public key μαζί με το αντίγραφο στην στοίβα.
3. Ο τελεστής εντολής OP_HASH160:
 - > Αφαιρεί το αντίγραφο του public key του Bob από την στοίβα.
 - > Εφαρμόζει την συνάρτηση κατακερματισμού SHA256 πάνω στο αντίγραφο.
 - > Εφαρμόζει τον RIPEMD160 αλγόριθμο, ώστε να παραχθεί η public address του.
4. Ώθηση της public address του Bob στην στοίβα.

5. Ο τελεστής OP_EQUALVERIFY:

- > Αφαιρεί τα δύο τελευταία στοιχεία από την κορυφή της στοίβας.
- > Ελέγχει αν οι public addresses είναι έγκυρες ή όχι. Στην περίπτωση που δεν είναι έγκυρες σταματά και η διαδικασία ακυρώνεται.

6. Ο τελεστής OP_CHECKSIG:

- > Αφαιρεί από την στοίβα το public key και την signature του Bob.
- > Ελέγχει αν η signature και το Public key είναι έγκυρα.

Όταν ολοκληρωθεί με επιτυχία, το αποτέλεσμα της στοίβας έχει την τιμή 1.

Στις εικόνες 14 και 15 αναλύεται το παράδειγμα 2:

<Bob's signature> <Bob's public key> OP_DUP OP_HASH160 <Bob's public address>
OP_EQUALVERIFY OP_CHECKSIG.

Step 1:

<Bob's signature> <Bob's public key> OP_DUP OP_HASH160 <Bob's public address>
OP_EQUALVERIFY OP_CHECKSIG

Pushing <Bob's signature> on to the stack

<Bob's signature>

<Bob's signature> <Bob's public key> OP_DUP OP_HASH160 <Bob's public address>
OP_EQUALVERIFY OP_CHECKSIG

Followed by <Bob's public key>

<Bob's public key>
<Bob's signature>

Step 2:

<Bob's signature> <Bob's public key> OP_DUP OP_HASH160 <Bob's public address>
OP_EQUALVERIFY OP_CHECKSIG

Now OP_DUP comes into play which first pops <Bob's public key> out of the stack.

<Bob's signature>

Duplicates it, and then pushes both the original and the duplicate onto the stack

Εικόνα 14: Παράδειγμα υλοποίησης P2PKH [13].

<Bob's public key>
<Bob's public key>
<Bob's signature>

Step 3:

<Bob's signature> <Bob's public key> OP_DUP OP_HASH160 <Bob's public address>
OP_EQUALVERIFY OP_CHECKSIG

The OP_HASH160 opcode pops out <Bob's public key> and runs it through the SHA 256 algorithm followed by RIPEMD160 algo to get <Bob's public address>

<Bob's public address>
<Bob's public key>
<Bob's signature>

Step 4:

So, what is the next element that gets pushed onto the stack?

<Bob's signature> <Bob's public key> OP_DUP OP_HASH160 <Bob's public address>
OP_EQUALVERIFY OP_CHECKSIG

<Bob's public address>
<Bob's public address>
<Bob's public key>
<Bob's signature>

Step 5:

<Bob's signature> <Bob's public key> OP_DUP OP_HASH160 <Bob's public address>
OP_EQUALVERIFY OP_CHECKSIG

Step 6:

<Bob's signature> <Bob's public key> OP_DUP OP_HASH160 <Bob's public address>
OP_EQUALVERIFY OP_CHECKSIG

Εικόνα 15: Συνέχεια παραδείγματος εικόνας 14 [13].

3.2.3.2 Pay to Public Key – P2PK

Η πληρωμή με δημόσιο κλειδί είναι ο πρώτος τύπος script που χρησιμοποιήθηκε για την διεκπεραίωση των συναλλαγών στο δίκτυο Bitcoin. Είναι πιο απλό από το P2PKH, καθώς αποθηκεύεται μόνο το public key. Χρησιμοποιείται κυρίως στις coinbase transactions, και οι εντολές σεναρίου είναι:

> scriptPubKey = <public key> OP_CHECKSIG

> scriptSig = <signature>

3.2.3.3 Pay to Script Hash – P2SK

Επειδή τα δεδομένα είναι σε μορφή διαδοχικών byte, εμφανίστηκε το 2012 ως μια λύση ώστε να χρησιμοποιούνται λιγότερα bytes για την αποθήκευση των δεδομένων. Οι εντολές σεναρίου είναι:

```
> scriptPubKey = OP_HASH160 <redeem script hash> OP_EQUAL
```

```
> scriptSig = <signature> <redeemerScript>
```

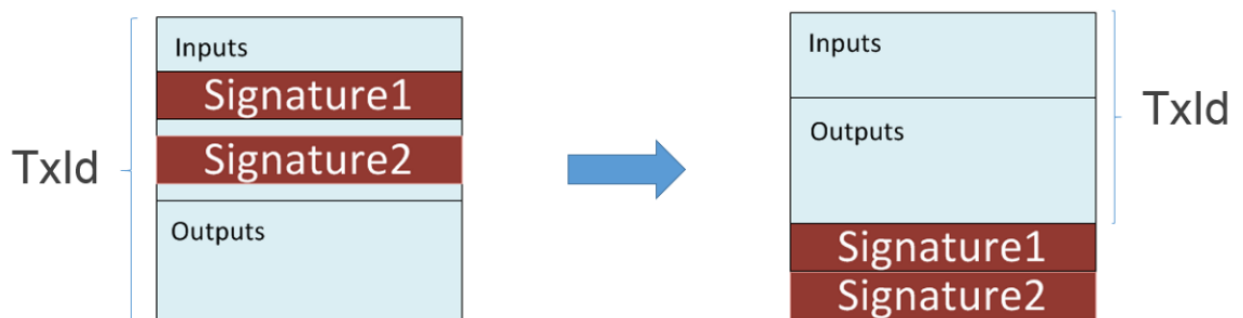
3.2.3.4 Pay to Witness Public Key Hash – P2WPKH

Μια σημαντική αναβάθμιση πρωτοκόλλου του δικτύου Bitcoin, είναι το P2WPKH. Λειτουργεί χωρίς ψηφιακή υπογραφή (Segregated Witness – SegWit) και στόχος είναι η αύξηση της χωρητικότητας ενός block (υπήρχε περιορισμός στο 1 MB). Τα δεδομένα των υπογραφών αφαιρούνται από το block βάσης, και μεταδίδονται στο δίκτυο μέσω ενός εξωτερικού block. Το block βάσης περιέχει πληροφορίες σχετικά με τον αποστολέα και τον παραλήπτη, ενώ το εξωτερικό block μεταδίδει τα δεδομένα υπογραφών. Κατά συνέπεια, με την αύξηση αποθηκευτικού χώρου οι συναλλαγές γίνονται πιο γρήγορα και επαληθεύονται πιο πολλές ανά block.

Οι εντολές σεναρίου είναι:

```
> scriptPubKey = OP_0 <public key hash>
```

```
> scriptSig = <signature> <public key>
```



Εικόνα 16: Διαφορά μεταξύ P2WPKH και P2PKH [15].

3.2.3.5 Pay to Witness Script Hash – P2WSH

Ισχύει ότι και στο σενάριο P2SH, δίνοντας την δυνατότητα αποθήκευσης περισσότερων χρήσιμων πληροφοριών των συναλλαγών.

Οι εντολές σεναρίου είναι:

```
> scriptPubKey = OP_0 <witness script hash> OP_EQUAL  
> scriptSig = <signature> <witnessScript>
```

3.2.3.6 Τελεστής επιστροφής (OP Return)

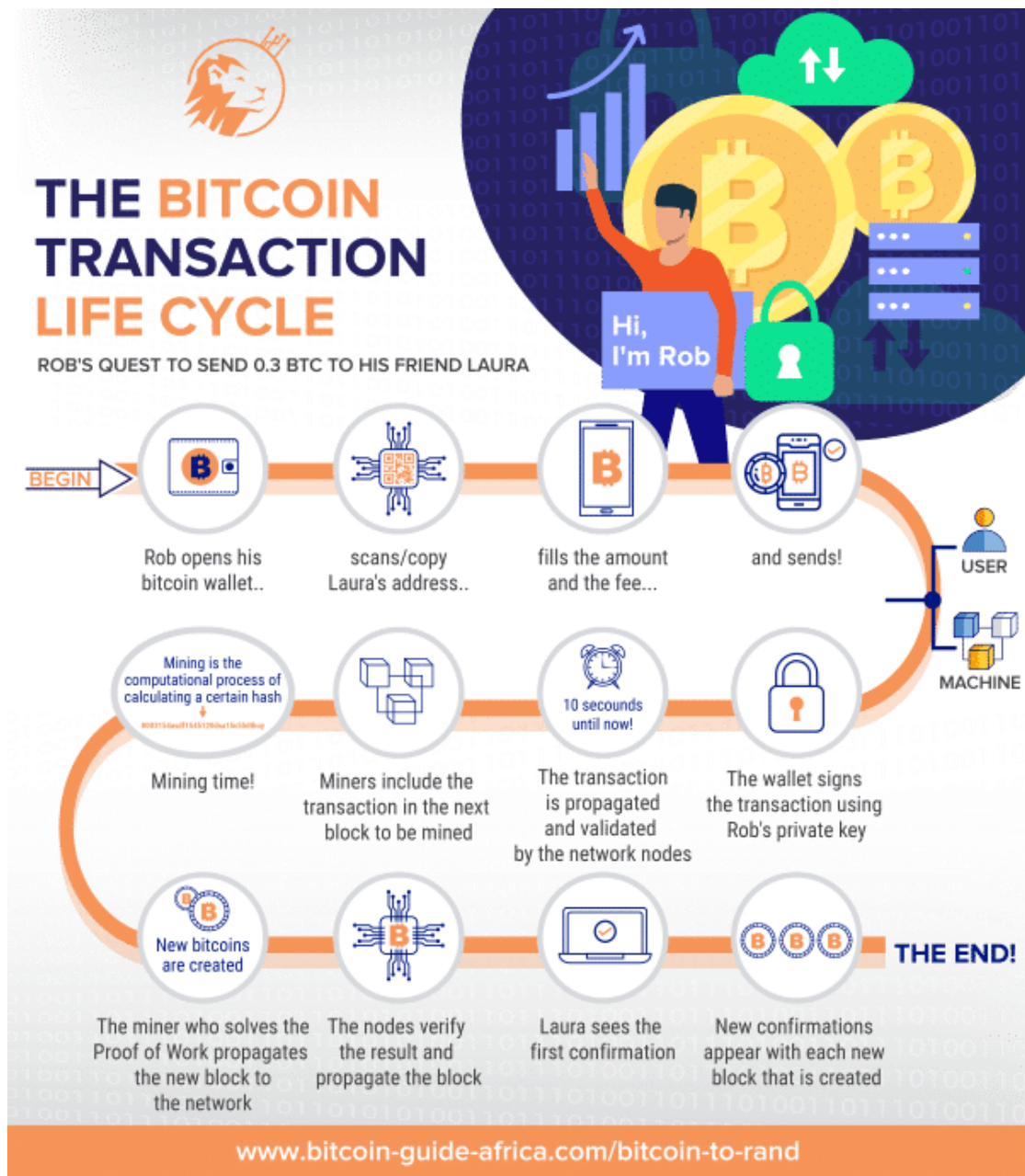
Ο τελεστής OP_RETURN χρησιμοποιείται για μη έγκυρες UTXO. Μόλις εκτελεστεί το script επισημαίνεται άκυρο και απορρίπτεται.

Οι εντολές σεναρίου είναι:

```
> scriptPubKey = OP_RETURN <data>  
> scriptSig = NULL
```

3.2.4 Εξόρυξη των blocks (Mining)

Το mining είναι μια διεργασία που χρησιμοποιείται για την επιβεβαίωση των εκκρεμών συναλλαγών και την αποθήκευσή τους στο Blockchain. Η διαδικασία αυτή προσφέρει ασφάλεια και λύνει το πρόβλημα της διπλής δαπάνης (double-spending), δηλαδή την δαπάνη της ίδια ποσότητας Bitcoin σε περισσότερες από μια συναλλαγές. Κάθε block εξορύσσεται περίπου κάθε 10 λεπτά και οι επιβεβαιωμένες από τους miners συναλλαγές προστίθενται στο Blockchain για την δαπάνη ως έγκυρες UTXO. Λόγω της αρχιτεκτονικής του δικτύου Bitcoin, κάθε block έχει περιορισμό σε χωρητικότητα 1MB. Κατά συνέπεια, οι συναλλαγές που πραγματοποιούνται αποθηκεύονται προσωρινά σε μια πισίνα μνήμης (memory pool) μέχρι να επαληθευτούν από τους miners. Οι miners ανταγωνίζονται μεταξύ τους για την επίλυση του παζλ κατακερματισμού και ο πρώτος που θα λύσει το παζλ, θα λάβει και μια ανταμοιβή.



Εικόνα 17: Κύκλος ζωής μιας συναλλαγής Bitcoin [16].

3.2.4.1 Κεφαλαιοποίηση Bitcoin

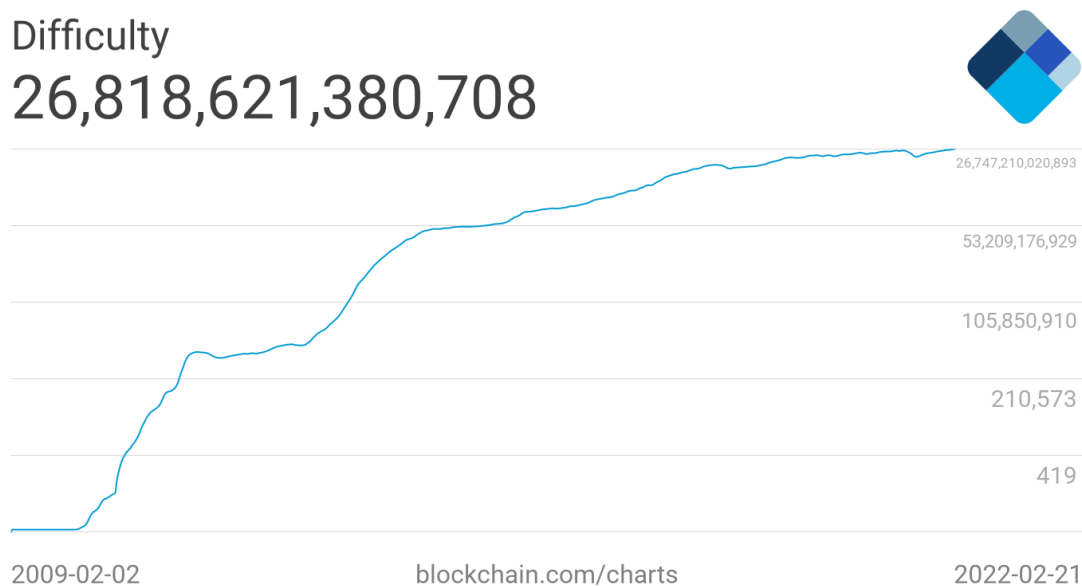
Τα Bitcoins παράγονται κάθε φορά που δημιουργείται ένα block, δηλαδή κάθε 10 λεπτά. Η ανταμοιβή υποδιπλασιάζεται με σταθερό ρυθμό κάθε 210.000 block ή κάθε 4 χρόνια (κατά μέσο όρο). Υπολογίζεται ότι το ανώτατο όριο θα είναι 21.000.000 Bitcoin και θα έχουν παραχθεί μέχρι το έτος 2078. Ωστόσο, θα συνεχιστούν να παράγονται blocks αλλά οι miners θα λαμβάνουν ως ανταμοιβή τον φόρο συναλλαγής (transaction fees). Το 2009 η ανταμοιβή ήταν 50 BTC, ενώ σήμερα το 2022 η ανταμοιβή είναι 6.25 BTC.

3.2.4.2 Παζλ κατακερματισμού (hash)

Οι miners πρέπει να βρουν έναν 64 ψηφίο δεκαεξαδικό αριθμό, δηλαδή ένα κατακερματισμό (hash) το οποίο να είναι μικρότερο ή ίσο από τον κατακερματισμό στόχου (target hash). Το target hash είναι μια πληροφορία στο κάθε block και καθορίζει τον βαθμό δυσκολίας του mining. Οι miners δοκιμάζουν διαφορετικούς συνδυασμούς hash μέχρι να καταλήξουν σε μια λύση. Η λύση που είναι μικρότερη ή ίση με το target hash λαμβάνει την ανταμοιβή.

3.2.4.3 Βαθμός δυσκολίας

Ο βαθμός δυσκολίας του mining αναφέρεται στο πόσο δύσκολο είναι να βρεθεί το target hash και συνεπώς επηρεάζει τον ρυθμό που παράγονται τα Bitcoin. Ο Satoshi έχει ορίσει την δυσκολία αυτή να αλλάζει κάθε 2.016 block, καθώς ο ρυθμός που παράγονται τα blocks πρέπει να είναι σταθερός.



Εικόνα 18: Βαθμός δυσκολίας σε βάθος χρόνου (λογαριθμική κλίμακα) [17].

Κεφάλαιο 4. Σχεδίαση & Υλοποίηση

Σε αυτό το κεφάλαιο περιγράφεται λεπτομερώς η διαδικασία της αποκωδικοποίησης των συναλλαγών. Αναλύονται τα βήματα εγκατάστασης του δικτύου Bitcoin, τα περιεχόμενα κάθε block καθώς και η σύνδεση κάθε block συναλλαγών με το προηγούμενο. Επίσης περιγράφονται οι μέθοδοι για την ομαδοποίηση των διευθύνσεων που ανήκουν στον ίδιο χρήστη και η εκ νέου ταυτοποίηση, καθώς και ο τρόπος λειτουργίας του αποκωδικοποιητή.

4.1 Εγκατάσταση Bitcoin Core – Full Node

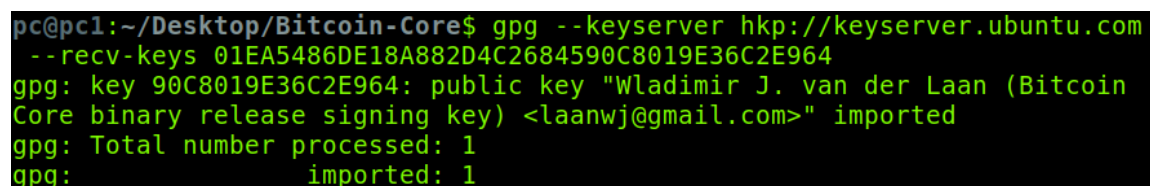
4.1.1 Επαλήθευση λήψης του Bitcoin Core

Απαραίτητο βήμα για την εγκατάσταση και τον συγχρονισμό του Bitcoin Core, είναι η επαλήθευση της έκδοσής του (release), μέσω των “υπογραφών” (signatures). Οι εκδόσεις 0.11, αλλά και οι νεότερες εκδόσεις, του Bitcoin Core υπογράφονται από το κλειδί του Wladimir J. van der Laan με το δακτυλικό αποτύπωμα:

01EA5486DE18A882D4C2684590C8019E36C2E964.

Το αρχείο που περιέχει την υπογραφή για την επαλήθευση της τρέχουσας έκδοσης (verify release signatures) είναι το [SHA256SUMS.asc](#). Στον φάκελο που βρίσκεται το αρχείο λήψης [SHA256SUMS.asc](#) γίνεται επαλήθευση χρησιμοποιώντας τις παρακάτω εντολές:

```
$ gpg --keyserver hkp://keyserver.ubuntu.com  
--recv-keys 01EA5486DE18A882D4C2684590C8019E36C2E964
```



```
pc@pc1:~/Desktop/Bitcoin-Core$ gpg --keyserver hkp://keyserver.ubuntu.com  
--recv-keys 01EA5486DE18A882D4C2684590C8019E36C2E964  
gpg: key 90C8019E36C2E964: public key "Wladimir J. van der Laan (Bitcoin  
Core binary release signing key) <laanwj@gmail.com>" imported  
gpg: Total number processed: 1  
gpg: imported: 1
```

Εικόνα 19: Επαλήθευση των υπογραφών.

```
$ gpg --verify SHA256SUMS.asc
```

```
pc@pcl:~/Desktop/Bitcoin-Core$ gpg --verify SHA256SUMS.asc
gpg: Signature made Thu 14 Jan 2021 02:35:31 PM EET
gpg: using RSA key 90C8019E36C2E964
gpg: Good signature from "Wladimir J. van der Laan (Bitcoin Core binary release signing key) <laanwj@gmail.com>" [unknown]
gpg: WARNING: This key is not certified with a trusted signature!
gpg: There is no indication that the signature belongs to the owner.
Primary key fingerprint: 01EA 5486 DE18 A882 D4C2 6845 90C8 019E 36C2 E964
```

Εικόνα 20: Καλή υπογραφή - Έγκυρο Bitcoin Core.

Η “καλή υπογραφή” (good signature) σημαίνει ότι η υπογραφή επαληθεύτηκε και όλα είναι εντάξει για την εγκατάσταση του Bitcoin Core.

4.1.2 Εγκατάσταση Bitcoin Core

Το αρχείο εγκατάστασης του [Bitcoin Core](#) (σε συμπιεσμένη μορφή .tar.gz), είναι απαραίτητο να αποθηκευτεί στον ίδιο φάκελο που έγινε η λήψη του αρχείου [SHA256SUMS.asc](#). Μεταβαίνοντας στον φάκελο αυτόν (έστω ο φάκελος είναι ~/Desktop/Bitcoin_Core/), οι εντολές για την εγκατάσταση είναι οι παρακάτω:

```
$ tar xzf bitcoin-0.21.0-x86_64-linux-gnu.tar.gz,
```

- όπου θα δημιουργηθεί ένας νέος φάκελος bitcoin-0.21.0

Στην συνέχεια, μεταβαίνουμε στο path ~/Desktop/Bitcoin-Core/bitcoin-0.21.0/bin/ και εκτελούμε:

```
$ sudo install -m 0755 -o root -g root -t /usr/local/bin *
```

Η παραπάνω εντολή θα μετακινήσει τα αρχεία

```
bitcoin-cli bitcoind bitcoin-qt bitcoin-tx bitcoin-wallet test_bitcoin
```

στο /usr/local/bin/ ενώ τους δίνει τα κατάλληλα δικαιώματα που χρειάζονται. Τέλος, η εγκατάσταση θα γίνει μέσω του Bitcoin Core Graphical User Interface (GUI), με την εντολή:

```
$ /usr/local/bin/bitcoin-qt
```

```
pc@pcl:~$ /usr/local/bin/bitcoin-qt
█
```

Εικόνα 21: Εντολή έναρξης του οδηγού εγκατάστασης.

Για να λειτουργήσει σωστά το Bitcoin Core GUI, χρειάζονται να εγκατασταθούν ορισμένες βιβλιοθήκες (libraries), οπότε σε περίπτωση εμφάνισης ενός μηνύματος με σφάλμα π.χ.

/usr/local/bin/bitcoin-qt: error while loading shared libraries:

libQtGui.so.4: cannot open shared object file: No such file or directory

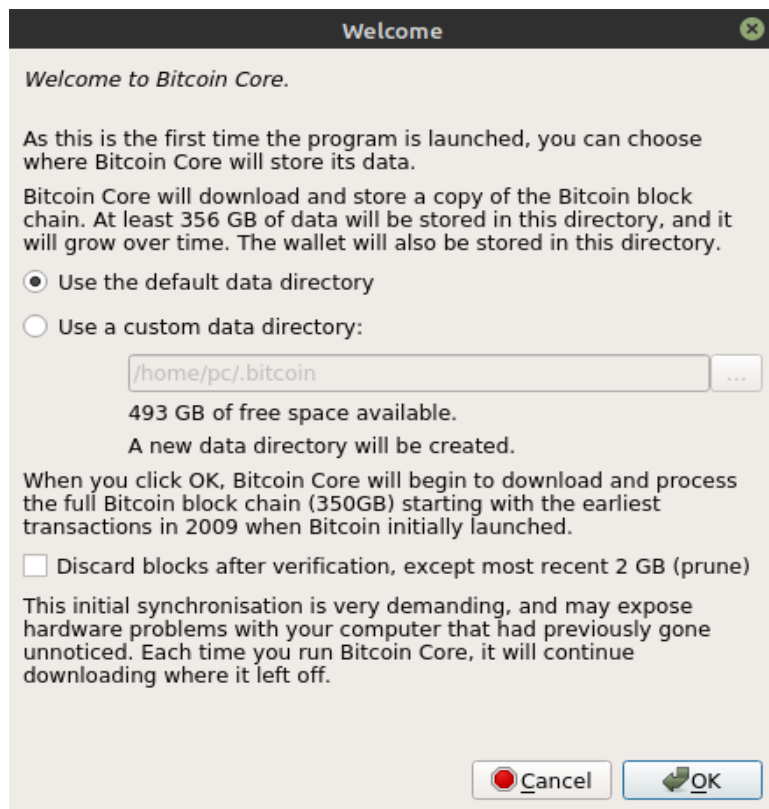
πρέπει να εγκατασταθούν οι βιβλιοθήκες που λείπουν. Μετά την εγκατάσταση των libraries (εάν χρειαστεί), εκτελείτε πάλι η εντολή:

\$ /usr/local/bin/bitcoin-qt.



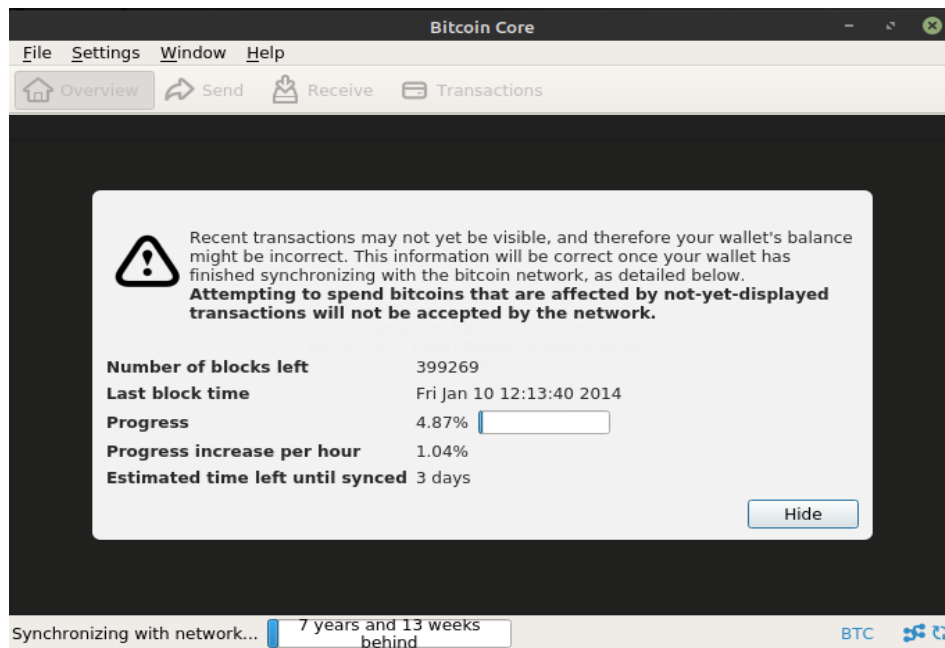
Εικόνα 22: Οδηγός εγκατάστασης - Setup Wizard.

Αφού η διαδικασία ολοκληρωθεί σωστά, το Bitcoin Core GUI είναι έτοιμο να ξεκινήσει.



Εικόνα 23: Επιλογή φακέλου αποθήκευσης δεδομένων.

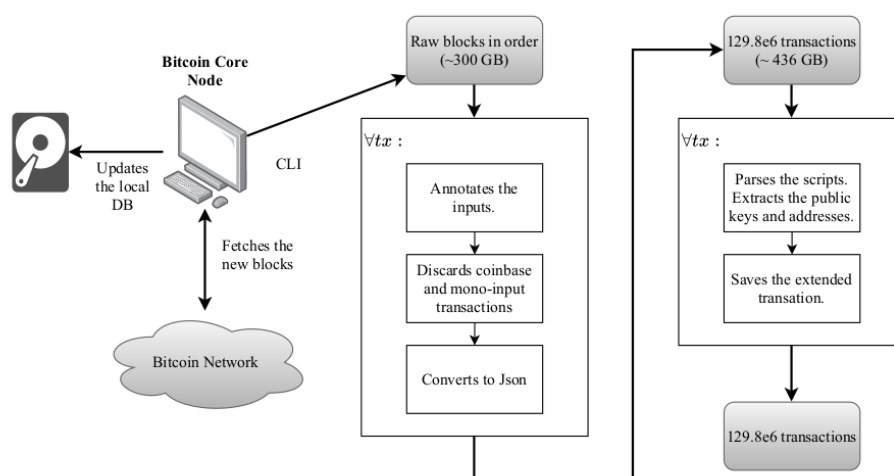
Λόγω του μεγάλου όγκου των δεδομένων θα χρειαστεί 3-7 μέρες για τον συγχρονισμό του Bitcoin Core.



Εικόνα 24: Συγχρονισμός των block συναλλαγών.

4.2 Αποκωδικοποίηση δεδομένων

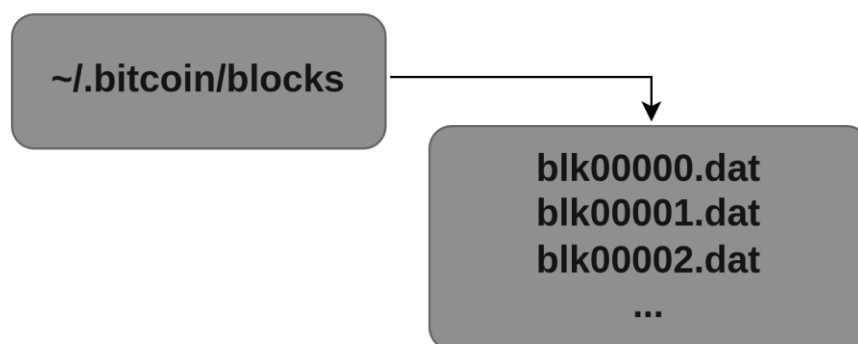
Τα δεδομένα είναι αποθηκευμένα στο Bitcoin Core, δηλαδή σε ένα πλήρη κόμβο (full node). Ένας πλήρης κόμβος συγχρονίζεται άμεσα με τους άλλους πλήρεις κόμβους του δικτύου, οι οποίοι αποδέχονται και επικυρώνουν τις συναλλαγές και τα blocks. Με αυτόν τον τρόπο το δίκτυο ενημερώνεται αυτόματα με την δημιουργία κάθε καινούργιου block, ενώ το καινούργιο block αποθηκεύεται στον σκληρό δίσκο.



Εικόνα 25: Ροή δεδομένων [18].

Επίσης, αντί να αποθηκεύεται ολόκληρο το Blockchain σε ένα αρχείο με μεγάλη χωρητικότητα μνήμης, χωρίζεται σε πολλά μικρά αρχεία με την κατάληξη της μορφής blk*.dat (* = ένας πενταψήφιος αριθμός ξεκινώντας από το μηδέν, δηλαδή blk00000.dat). Ο λόγος είναι ότι το μέγιστο μέγεθος κάθε αρχείου είναι 128MiB (134.217.728 bytes). Κατά την διάρκεια του συγχρονισμού κάθε νέο block προστίθεται στο τέλος του τρέχων αρχείου και όταν αυτό φτάσει στην επιτρεπόμενη μέγιστη χωρητικότητα τότε προχωρά στο επόμενο αρχείο και ούτω καθεξής. Μετά τον πλήρη συγχρονισμό (από προεπιλογή) τα αρχεία που περιέχουν τις συναλλαγές, βρίσκονται στον φάκελο με path:

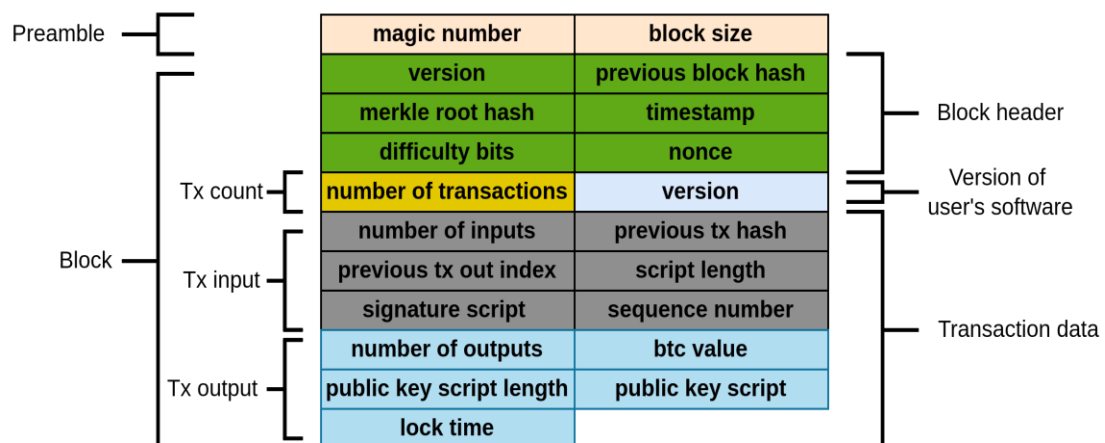
`~/home/[username]/.bitcoin/blocks`



Εικόνα 26: Το Blockchain χωρίζεται σε αρχεία μεγέθους 128MiB.

4.2.1 Δομή block

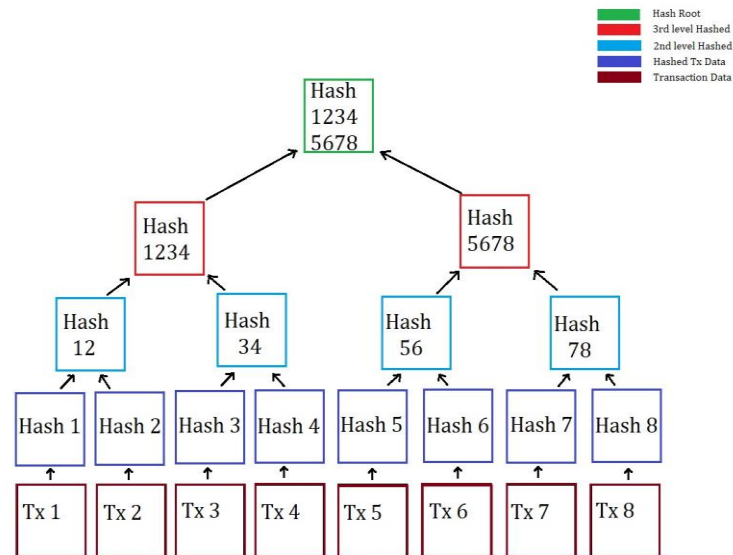
Τα δεδομένα συναλλαγών στα αρχεία blk*.dat αποθηκεύονται σε δυαδική μορφή.



Εικόνα 27: Δομή κάθε block.

Κάθε block περιέχει τα εξής πεδία:

- Μαγικός αριθμός (magic number): Είναι ένα αναγνωριστικό για το δίκτυο Blockchain. Έχει σταθερή τιμή 0xD9B4BEF9 και προσδιορίζει πότε ξεκινάει το κάθε block.
- Μέγεθος block (block size): Υποδεικνύει πόσο μεγάλο είναι το block. Μαζί με τον μαγικό αριθμό προσδιορίζουν πότε ξεκινάει και πότε τελειώνει το κάθε block.
- Έκδοση (version): Η έκδοση του λογισμικού Bitcoin.
- Δέντρο merkle ή ρίζα κατακερματισμού merkle (merkle root hash ή merkle tree): Το δέντρο merkle περιέχει τις συναλλαγές σε ζεύγη μετά την διαδικασία κατακερματισμού. Οι κατακερματισμοί που προκύπτουν κατακερματίζονται σε ζεύγη μέχρι να παραμείνει μόνο ένας τελικός κατακερματισμός (merkle root hash).



Εικόνα 28: Δέντρο merkle ή ρίζα κατακερματισμού merkle [19].

- Χρονοσφαιίδα (timestamp): Είναι η χρονική στιγμή που δημιουργήθηκε το block, σε μορφή Unix Timestamp. Το Unix Timestamp είναι ο αριθμός των δευτερολέπτων που έχουν περάσει από την 01.01.1970, που σημαίνει ότι 0000000000 σε χρόνο UNIX είναι ίσο με την 1η Ιανουαρίου 1970, 12:00:00 π.μ.
- Bits δυσκολίας (difficulty bits): Αριθμητική αναπαράσταση της τρέχουσας δυσκολίας.
- Αριθμός που χρησιμοποιείται μόνο μία φορά (nonce): Είναι σημαντικό στην κρυπτογραφία για την απόδειξη της εργασίας (PoW). Αναφέρεται στον αριθμό που πρέπει να βρει ο miner για την λύση του παζλ.
- Αριθμός συναλλαγών (tx count): Είναι ο αριθμός των συναλλαγών που περιλαμβάνονται στο block.

- [illegible]

27

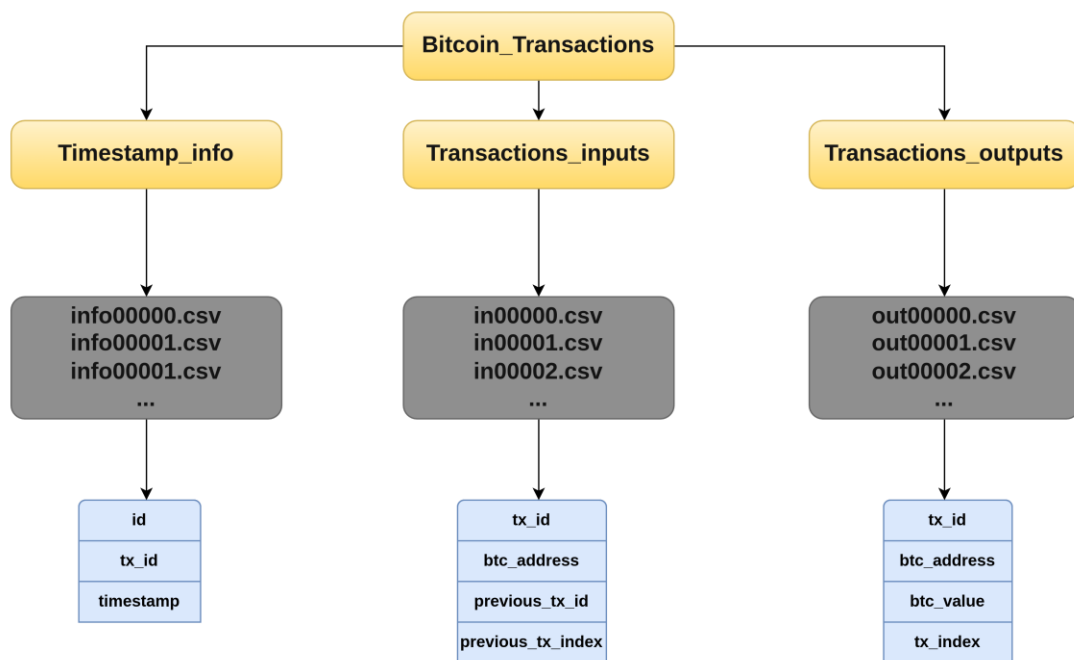
Για παράδειγμα, διαβάζοντας byte-byte το genesis μπλοκ, τα πεδία είναι (σε παρένθεση τα bytes που απαιτούνται):

- magic number (4 bytes): f9be b4d9
- block size (4 bytes): 1d01 0000
- version (4 bytes): 0100 0000
- previous block hash (32 bytes): 0000 0000 0000 0000 0000 0000 0000 0000
0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000
- merkle root (32 bytes): 3ba3 edfd 7a7b 12b2 7ac7 2c3e 6776 8f61 7fc8 1bc3 888a
5132 3a9f b8aa 4b1e 5e4a
- timestamp (4 bytes): ffff 001d
- difficulty bits (4 bytes): 29ab 5f49
- nonce (4 bytes): 1dac 2b7c
- number of transactions (1-9 bytes): 01
- version of user's software (4 bytes): 0100 0000
- number of inputs (1-9 bytes): 01
- previous tx hash (32 bytes): 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000
0000 0000 0000 0000 0000 0000 0000 0000
- previous tx index (4 bytes): ffff ffff
- script length (1-9 bytes): 4d
- signature script (varies): 04 ffff 001d 0104 4554 6865 2054 696d 6573 2030 332f
4a61 6e2f 3230 3039 2043 6861 6e63 656c 6c6f 7220 6f6e 2062 7269 6e6b 206f
6620 7365 636f 6e64 2062 6169 6c6f 7574 2066 6f72 2062 616e 6b73
- sequence number (4 bytes): ffff ffff
- number of outputs (1-9 bytes): 01
- btc value (8 bytes): 00f2 052a 0100 0000
- public key script length (1-9 bytes): 43
- public key script (varies): 4104 678a fdb0 fe55 4827 1967 f1a6 7130 b710 5cd6
a828 e039 09a6 7962 e0ea 1f61 deb6 49f6 bc3f 4cef 38c4 f355 04e5 1ec1 12de
5c38 4df7 ba0b 8d57 8a4c 702b 6bf1 1d5f ac
- locktime (4 bytes): 0000 0000

4.2.2 Δεδομένα συναλλαγών

Κατά την αποκωδικοποίηση των δεδομένων δημιουργείται ένας φάκελος ο οποίος περιέχει τρεις φακέλους (έστω Bitcoin_Transactions).

Για κάθε αρχείο blk*.dat θα δημιουργηθεί ένα αντίστοιχο αρχείο σε κάθε φάκελο, όπου περιέχουν διαφορετικές πληροφορίες για τις συναλλαγές.



Εικόνα 30: Περιεχόμενα καταλόγου Bitcoin_Transactions.

Αυτό γίνεται γιατί οι διευθύνσεις εισόδου δεν δίνονται άμεσα, αλλά πρέπει να βρεθούν με βάση το αναγνωριστικό (previous_tx_id) και τον δείκτη της προηγούμενης συναλλαγής (previous_tx_index).

Ο φάκελος Timestamp_info περιέχει αρχεία με τις εξής πληροφορίες για τα αναγνωριστικά των συναλλαγών:

- id = μοναδικό αναγνωριστικό για το σύνολο των συναλλαγών
- tx_id = αναγνωριστικό συναλλαγής
- timestamp = ώρα και ημερομηνία που έγινε η συναλλαγή σε unix timestamp μορφή

Πίνακας 1: Παράδειγμα πεδίων του αρχείου info00000.csv.

id	tx_id	timestamp
1	4a5e1e4baab89f3a32518a88c31bc87f618f76673e2cc77ab2127b7afdeda33b	1231006505
2	0e3e2357e806b6cdb1f70b54c3a3a17b6714ee1f0e68bebb44a74b1efd512098	1231469665
3	9b0fc92260312ce44e74ef369f5c66bbb85848f2eddd5a7a1cde251e54ccfdd5	1231469744
4	999e1c837c76a1b7fbb7e57baf87b309960f5ffefbf2a9b95dd890602272f644	1231470173
5	df2b060fa2e5e9c8ed5eaf6a45c13753ec8c63282b2688322eba40cd98ea067a	1231470988
...

Ο φάκελος Transaction_inputs περιέχει αρχεία με πληροφορίες για τις διευθύνσεις εισόδου:

- tx_id = αναγνωριστικό συναλλαγής
- btc_address = διεύθυνση αποστολέα
- previous_tx_id = αναγνωριστικό προηγούμενης συναλλαγής
- previous_tx_index = δείκτης προηγούμενης συναλλαγής

Πίνακας 2: Παράδειγμα πεδίων για διευθύνσεις εισόδου του αρχείου in00000.csv.

tx_id	btc_address	previous_tx_id	previous_tx_index
4a5e1e4baab89f3a32518a88c31bc87f618f76673e2cc77ab2127b7afdeda33b	MINING_REWARD	0x	-1
0e3e2357e806b6cdb1f70b54c3a3a17b6714ee1f0e68bebb44a74b1efd512098	MINING_REWARD	0x	-1
9b0fc92260312ce44e74ef369f5c66bbb85848f2eddd5a7a1cde251e54ccfdd5	MINING_REWARD	0x	-1
999e1c837c76a1b7fbb7e57baf87b309960f5ffefbf2a9b95dd890602272f644	MINING_REWARD	0x	-1
df2b060fa2e5e9c8ed5eaf6a45c13753ec8c63282b2688322eba40cd98ea067a	MINING_REWARD	0x	-1
...

*Σημείωση: Για την δημιουργία μπλοκ, σε αντίθεση με την μεταφορά bitcoin, το πεδίο pre-vius tx hash παίρνει την τιμή 0x και το πεδίο previous tx index παίρνει την τιμή -1.

Ο φάκελος Transaction_outputs περιέχει αρχεία με πληροφορίες για τις διευθύνσεις εξόδου:

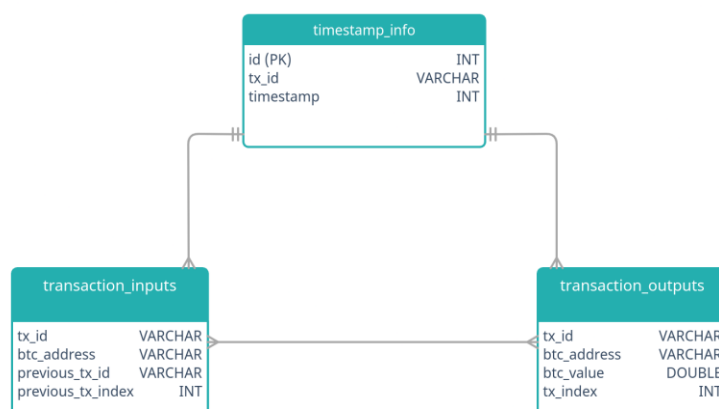
- tx_id = αναγνωριστικό συναλλαγής
- btc_address = διεύθυνση αποστολέα/παραλήπτη
- btc_value= ποσό συναλλαγής σε αξία
- bitcoin tx_index = δείκτης τρέχουσας συναλλαγής

Πίνακας 3: Παράδειγμα πεδίων για διευθύνσεις εξόδου του αρχείου out00000.csv.

tx_id	btc_address	btc_value	tx_index
4a5e1e4baab89f3a32518a88c31bc87f618f76673e2cc77ab2127b7afdeda33b	1A1zP1eP5QGefi2DMPTfTL5SLmv7DivfNa	50	0
0e3e2357e806b6cdb1f70b54c3a3a17b6714ee1f0e68bebb44a74b1efd512098	12c6DSiU4Rq3P4ZxziKxziL5LmMBRzjrjX	50	0
9b0fc92260312ce44e74ef369f5c66bbb85848f2edd5a7a1cde251e54ccfdd5	1HLoD9E4SDFFPDiYfNynkBLQ85Y51J3Zb1	50	0
999e1c837c76a1b7fbb7e57baf87b309960f5ffefbf2a9b95dd890602272f644	1FvzCLoTPGANNjWoUo6jUGuAG3wg1w4YjR	50	0
df2b060fa2e5e9c8ed5eaf6a45c13753ec8c63282b2688322eba40cd98ea067a	15ubicBBWFnvoZLT7GiU2qxjRaKJPdkDMG	50	0
...

4.2.3 Άγνωστες διευθύνσεις εισόδου

Οι άγνωστες διευθύνσεις εισόδου, θα βρεθούν από τις πληροφορίες των εξερχόμενων συναλλαγών, με βάση το αναγνωριστικό και τον δείκτη της προηγούμενης συναλλαγής.



Εικόνα 31: Απεικόνιση των tables της βάσης δεδομένων και της σχέσης μεταξύ τους.

Για την γρήγορη και αποτελεσματική επεξεργασία των δεδομένων χρησιμοποιούνται βάσεις δεδομένων. Υπάρχουν 3 tables, όπου το καθένα αντιστοιχεί σε έναν φάκελο και περιέχει τα απαραίτητα πεδία.

Στην πρώτη φάση, θα ενημερωθούν οι άγνωστες διευθύνσεις εισόδου.

Για παράδειγμα,

- για την συναλλαγή με

tx_id_1: a16f3ce4dd5deb92d98ef5cf8afeaf0775ebca408f708b2146c4fb42b41e14be

οι πληροφορίες για τις διευθύνσεις εισόδου είναι:

```
{btc_address: NO_ADDRESS,  
previous_tx_id:  
f4184fc596403b9d638783cf57adfe4c75c605f6356fbc91338530e9831e9e16,  
previous_tx_index: 1}
```

- Στην συναλλαγή με

tx_id_2: f4184fc596403b9d638783cf57adfe4c75c605f6356fbc91338530e9831e9e16

οι διευθύνσεις εξόδου του tx_id_2 είναι:

```
[{btc_value: 10.0,  
btc_address: 1Q2TWHE3GMdB6BZKafqwxXtWAWgFt5JvM3},  
{btc_value: 40.0,  
btc_address: 12cbQLTFMXRnSzktFkuoG3eHoMeFtpTu3S}]
```

Οπότε, η διεύθυνση εισόδου του **tx_id_1**, δίνεται από την διεύθυνση εξόδου του προηγούμενου αναγνωριστικού συναλλαγής και δείκτη (**tx_id_2**).

Άρα η διεύθυνση του **tx_id_1** είναι: **12cbQLTFMXRnSzktFkuoG3eHoMeFtpTu3S**.

Στην δεύτερη και τελική φάση, κάθε tx_id θα αντιστοιχισθεί με το μοναδικό του timestamp.

TRANSACTION RECEIPT



TRANSACTION IDENTIFIER:

[a16f3ce4dd5deb92d98ef5cf8afeaf0775ebca408f708b2146c4fb42b41e14be](#)

TRANSACTION TIMESTAMP:

2009-01-12 06:02 (UTC)

CONFIRMED TRANSACTION:

Included in block: #181 on the Bitcoin blockchain

SENDERS (INPUTS):

#	SENDER	VALUE (BTC)
0	12cbQLTFMXRnSzktFkuoG3eHoMeFtpTu3S	40.00000000
		TOTAL: 40.00000000 BTC

RECIPIENTS (OUTPUTS):

#	RECIPIENT	VALUE (BTC)
0	1DUDsfc23Dv9sPMEk5RsrtfzCw5ofi5sVW	10.00000000
1	12cbQLTFMXRnSzktFkuoG3eHoMeFtpTu3S	30.00000000
/	*** MINER FEE ***	**
→		TOTAL: 40.00000000 BTC

Εικόνα 32: Απόδειξη συναλλαγής για το tx_id_1 [20].

- το αναγνωριστικό συναλλαγής – tx_id είναι:
[a16f3ce4dd5deb92d98ef5cf8afeaf0775ebca408f708b2146c4fb42b41e14be](#)
- οι διευθύνσεις εισόδου (αποστολέας) – inputs είναι:
[12cbQLTFMXRnSzktFkuoG3eHoMeFtpTu3S]
- οι διευθύνσεις εξόδου & ποσό(παραλήπτης) – outputs είναι:
[{btc_value: 30.0, btc_address: 12cbQLTFMXRnSzktFkuoG3eHoMeFtpTu3S},
{btc_value: 10.0, btc_address: 1DUDsfc23Dv9sPMEk5RsrtfzCw5ofi5sVW}]
- η ώρα και ημερομηνία – timestamp είναι:
1231740133, 12/01/2009

4.2.4 Επαναπροσδιορισμός χρήστη

Παρόλο που το πρωτόκολλο Bitcoin προσφέρει ιδιωτικότητα, οι συναλλαγές είναι δημόσιες. Κατά συνέπεια, ένας χρήστης ο οποίος θέλει να περιορίσει την ιχνηλάτηση των συναλλαγών που πραγματοποιεί και να κρατήσει το απόρρητο αυτών, μπορεί να χρησιμοποιήσει όσες διαφορετικές διευθύνσεις θέλει. Ωστόσο, εφαρμόζονται κάποια υδριβικά (heuristics) [ReRM10] ώστε να ανιχνευθούν οι διευθύνσεις που πιθανώς ανήκουν στον ίδιο χρήστη (user entity). Τα υδριβικά που χρησιμοποιούνται είναι:

- Όλες οι διευθύνσεις που χρησιμοποιούνται ως είσοδο για την ίδια συναλλαγή ανήκουν στην ίδια οντότητα ελέγχου, που ονομάζεται χρήστης*.
- Αν υπάρχουν ακριβώς δύο (2) διευθύνσεις εξόδου $a1$ και $a2$, και η μια εμφανίζεται για πρώτη φορά (έστω $a1$) και η άλλη έχει εμφανιστεί ξανά (έστω $a2$), τότε η $a1$ θεωρείται η διεύθυνση αλλαγής (change address).
- Μία διεύθυνση θεωρείται change address, εάν ικανοποιεί τις παρακάτω ιδιότητες:
 - Η συναλλαγή δεν έχει πληροφορίες για την δημιουργία block.
 - Η διεύθυνση δεν είναι μεταξύ των διευθύνσεων εισόδου.
 - Είναι η μόνη διεύθυνση εξόδου που εμφανίζεται για πρώτη φορά.

*Σημείωση: Όλες οι διευθύνσεις εισόδου της ίδιας συναλλαγής μπορεί να μην ανήκουν στον ίδιο χρήστη. Ο λόγος είναι η “πισίνα” εξόρυξης (mining pool). Μια πισίνα ή ομάδα εξόρυξης αποτελείται από miners που συνδυάζουν τους υπολογιστικούς τους πόρους μέσω ενός κοινού δικτύου για να μεγιστοποιήσουν την πιθανότητα εύρεσης ενός μπλοκ. Εάν οι miners, ανταγωνιστούν επιτυχώς και βρουν την λύση του παζλ, τότε μοιράζονται την ανταμοιβή.

4.3 Χρήση αποκωδικοποιητή

Πριν την εκτέλεση του αποκωδικοποιητή, ο χρήστης θα πρέπει να εκτελέσει την εντολή

```
$ pip3 install -r requirements.txt
```

για να εγκατασταθούν κάποιες βιβλιοθήκες που είναι απαραίτητες.

Επίσης, επειδή η αντιστοίχιση των διευθύνσεων θα γίνει με βάσεις δεδομένων, όλες οι συναλλαγές θα αποθηκευτούν σε ένα αρχείο. Ο χρήστης θα πρέπει να εκτελέσει τις παρακάτω εντολές για να δώσει τα κατάλληλα δικαιώματα που χρειάζεται η MySQL για να γράψει στον φάκελο.

- φάκελος1 = όνομα του φακέλου που θέλει να αποθηκευτούν οι συναλλαγές (έστω Bitcoin_Transactions)

- Για την δημιουργία του φακέλου

```
$ sudo mkdir ~/Desktop/Bitcoin_Transactions/
```

- Για τα κατάλληλα και απαραίτητα δικαιώματα

```
$ sudo chown mysql:mysql ~/Desktop/Bitcoin_Transactions/
```

```
$ sudo chmod 777 ~/Desktop/Bitcoin_Transactions/
```

4.3.1 Εκτέλεση κώδικα για την αποκωδικοποίηση

Ο αποκωδικοποιητής εκτελείται από το path ~/Bitcoin_Blockchain_Parser/blk_parser με την εντολή:

```
$ python3 Main.py
```

Ο χρήστης θα ρωτηθεί:

1. το path όπου βρίσκονται τα blk*.dat αρχεία.
2. αν θέλει να κρατήσει ή όχι τα αρχεία blk*.dat μετά την αποκωδικοποίηση (οι επιλογές θα πρέπει να είναι είτε ναι είτε όχι [y/n])
3. το path όπου θα αποθηκευτούν οι πληροφορίες της συναλλαγής

Μετά την ολοκλήρωση της αποκωδικοποίησης των συναλλαγών, οι πληροφορίες για τις διευθύνσεις εισόδου ενημερώνονται και αντί να δείχνουν στο αναγνωριστικό και στο δείκτη της προηγούμενης συναλλαγής, πλέον είναι η διεύθυνση Bitcoin. Οι πληροφορίες των συναλλαγών είναι αποθηκευμένες σε μορφή json στο αρχείο **bitcoin_info.json**, στο επιλεγμένο path (π.χ. ~/Desktop/Bitcoin_Transactions/bitcoin_info.json).

```
#####
# Bitcoin Blockchain blkXXXXX.dat parser  #
#                                         #
#      Copyright (C) 2021 Dimitrios Zervas #
#####

> This script aimed to parse raw Bitcoin Blockchain Transactions
> Read 'Hints.txt' before running the script!
-----

                                Input directory path

> Please enter the full path where blkXXXXX.dat files are stored
> /home/ze/.bitcoin/blocks
> Valid input path -- /home/ze/.bitcoin/blocks/ > Number of files: 4
-----

> All blkXXXXX.dat files are 0.537 GB
> So you must be carefull while running this parser
> Because it needs (extra) ~ 2.0 GB of storage for preprocessing
> Make sure you have enough free space

> If you want to keep blkXXXXX.dat files after parsing      > Press 'y'
> If you want to delete blkXXXXX.dat files after parsing    > Press 'n'
-----

> Should the parser delete the files? [y/n]      > n
-----
```

Εικόνα 33: Ερώτηση για την διαγραφή των αρχείων blk*.dat.

```
                                Export directory path

> Path must be different than Input directory path...
> Please enter the full path where Bitcoin transaction files will be stored
> /home/ze/Desktop/Bitcoin_Transactions
> Valid export path -- /home/ze/Desktop/Bitcoin_Transactions/
-----

on 0: blk00000.dat
on 1: blk00001.dat
on 2: blk00002.dat
on 3: blk00003.dat
|████████████████████████████████████████████████████████████████████████████████| 4/4 [100%] in 2:45.3 (0.02/s)
-----
```

Εικόνα 34: Ερώτηση για το path εξαγωγής των δεδομένων.

```

> Connection to Database Established...
> Database 'Bitcoin_Transactions' created successfully...
> Creating table: transaction_inputs
> Creating table: transaction_outputs
> Creating table: timestamp_info

> Loading Transaction_inputs into      - Database:      Bitcoin_Transactions
                                     - Table:          transaction_inputs

on 0: in000000.csv
on 1: in000001.csv
on 2: in000002.csv
on 3: in000003.csv
|████████████████████████████████████████| 4/4 [100%] in 41.7s (0.10/s)

> Done...
-----

> Loading Transaction_outputs into      - Database:      Bitcoin_Transactions
                                     - Table:          transaction_outputs

on 0: out000000.csv
on 1: out000001.csv
on 2: out000002.csv
on 3: out000003.csv
|████████████████████████████████████████| 4/4 [100%] in 55.1s (0.07/s)

> Done...
-----

> Loading Timestamp_info into          - Database:      Bitcoin_Transactions
                                     - Table:          timestamp_info

on 0: info000000.csv
on 1: info000001.csv
on 2: info000002.csv
on 3: info000003.csv
|████████████████████████████████████████| 4/4 [100%] in 17.9s (0.22/s)

> Done...
-----

```

Εικόνα 35: Πέρασμα αρχείων στην βάση δεδομένων.

```

> Readdressing
> This may take a while, please wait...
> Creating a new table with updated transaction inputs info
> Creating a new table with updated transaction outputs info
> Matching and cleaning up the data
> Saving data into folder '/home/ze/Desktop/Bitcoin_Transactions/'

> Done...
-----

> Bitcoin transactions are in: '/home/ze/Desktop/Bitcoin_Transactions/bitcoin_info.json'
> Don't forget to change file permissions, execute:
    $ sudo chmod 777 /home/ze/Desktop/Bitcoin_Transactions/bitcoin_info.json

> Exiting...
-----

```

Εικόνα 36: Ολοκλήρωσης εύρεσης διευθύνσεων και εξαγωγή των δεδομένων στο επιλεγμένο path.

Επίσης, λόγω του ότι η MySQL έχει αυστηρούς περιορισμούς σχετικά με το ποιος μπορεί να διαβάσει/γράψει σε ένα αρχείο που δημιουργείται από την εκτέλεση των εντολών της, ο χρήστης θα πρέπει να τρέξει την εντολή:

```
$ sudo chmod 777 ~/Desktop/Bitcoin_Transactions/bitcoin_info.json
```

για να έχει πρόσβαση στα τελικά δεδομένα. Τα τελικά δεδομένα είναι της μορφής:

```
{
    αναγνωριστικό συναλλαγής,
    διευθύνσεις εισόδου,
    διευθύνσεις εξόδου & ποσό,
    ώρα & ημερομηνία
}
```

Για παράδειγμα:

```
{
  tx_id:    f4184fc596403b9d638783cf57adfe4c75c605f6356fbc91338530e9831e9e16
  inputs:   [12cbQLTFMXRnSzktFkuoG3eHoMeFtpTu3S],
  outputs:  [{btc_value: 10.0, btc_address: 1Q2TWHE3GMdB6BZKafqwxXtWAWgFt5Jv...},
             {btc_value: 40.0, btc_address: 12cbQLTFMXRnSzktFkuoG3eHoMeFtpTu3S}],
  timestamp: 1231731025
}
```

Fee	0.00000000 BTC (0.000 sat/B - 0.000 sat/WU - 275 bytes)	50.00000000 BTC
Hash	f4184fc596403b9d638783cf57adfe4c75c605f6356fbc9133...	2009-01-12 05:30
	12cbQLTFMXRnSzktFkuoG3eHoMeFtpTu3S 50.00000000 BTC →	1Q2TWHE3GMdB6BZKafqwxXtWAWgFt5Jv... 10.00000000 BTC 12cbQLTFMXRnSzktFkuoG3eHoMeFtpTu3S 40.00000000 BTC

Εικόνα 37: Παράδειγμα εύρεσης διευθύνσεων [21].

4.3.2 Ομαδοποίηση των διευθύνσεων

Η εκτέλεση του κώδικα πραγματοποιείται από το path:

~/Bitcoin_Blockchain_Parser/community_detection, με την εντολή:

```
$python3 Track_Users.py file
```

όπου file, ολόκληρο το path που βρίσκεται το αρχείο bitcoin_info.json.

Στην αρχή, σύμφωνα με τα υδριβικά (Βλ. 4.2.4) ομαδοποιούνται οι διευθύνσεις που χρησιμοποιούνται ως είσοδο (εισροές) στην ίδια συναλλαγή καθώς και οι change διευθύν-

σεις. Αυτό έχει ως αποτέλεσμα την ιχνηλάτηση των διαφορετικών διευθύνσεων που ανήκουν στον ίδιο χρήστη. Στην συνέχεια, ενώ βρεθούν τα user entities, πραγματοποιείται ανάθεση μοναδικού αναγνωριστικού σε κάθε χρήστη ξεκινώντας από το μηδέν (0). Δημιουργείται σε μορφή txt ένα νέο αρχείο **reidentified_users.txt**, το οποίο αποθηκεύεται στον ίδιο φάκελο που βρίσκεται και το αρχείο bitcoin_info.json, δηλαδή στο path `~/Desktop/Bitcoin Transactions/`.

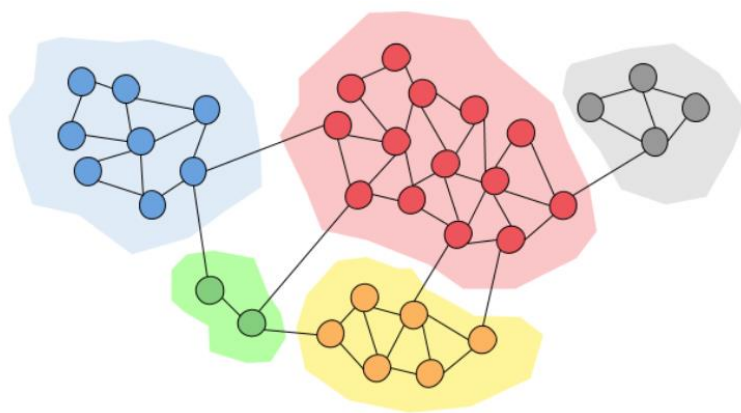
```
> Start processing  
> This may take a while, please wait...  
  
> Detecting same controlling entities & change addresses  
██████████████████████████████████████| 29995059/29995059 [100%] in 5:39.7 (88294.71/s)  
  
> Re-identifying users  
██████████████████████████████████████| 29995059/29995059 [100%] in 10:39.0 (46938.09/s)  
  
> Done...  
-----  
> Changes saved to '/home/ze/Desktop/Bitcoin_Transactions/reidentified_users.txt'  
> Exiting ...
```

Για παράδειγμα, για την συναλλαγή:

Η διεύθυνση 12cbQLTFMXRnSzktFkuoG3eHoMeFtpTu3S ανήκει στον χρήστη_0, η διεύθυνση 1Q2TWHE3GMdB6BZKafqwxXtWAWgFt5Jym3 ανήκει στον χρήστη_1, επομένως το τελικό αποτέλεσμα είναι: 0110.01231731025

Κεφάλαιο 5. Ανίχνευση κοινοτήτων

Τα δίκτυα αποτελούν μέσο αναπαράστασης συνδέσεων μεταξύ πολύπλοκων συστημάτων με μεγάλο όγκων δεδομένων, όπως για παράδειγμα στην περίπτωση της τεχνολογίας και της επικοινωνίας. Σχηματίζονται δίκτυα τεράστιας κλίμακας, τα οποία αποτελούνται από διάφορες κοινότητες. Πιο συγκεκριμένα, μια κοινότητα αναφέρεται στις ομάδες κόμβων που αλληλοεπιδρούν πιο συχνά μεταξύ τους σε σύγκριση με το υπόλοιπο δίκτυο, όπως φαίνεται στην παρακάτω εικόνα (το δίκτυο αποτελείται από 5 κοινότητες).



Εικόνα 39: Απεικόνιση κοινοτήτων [22].

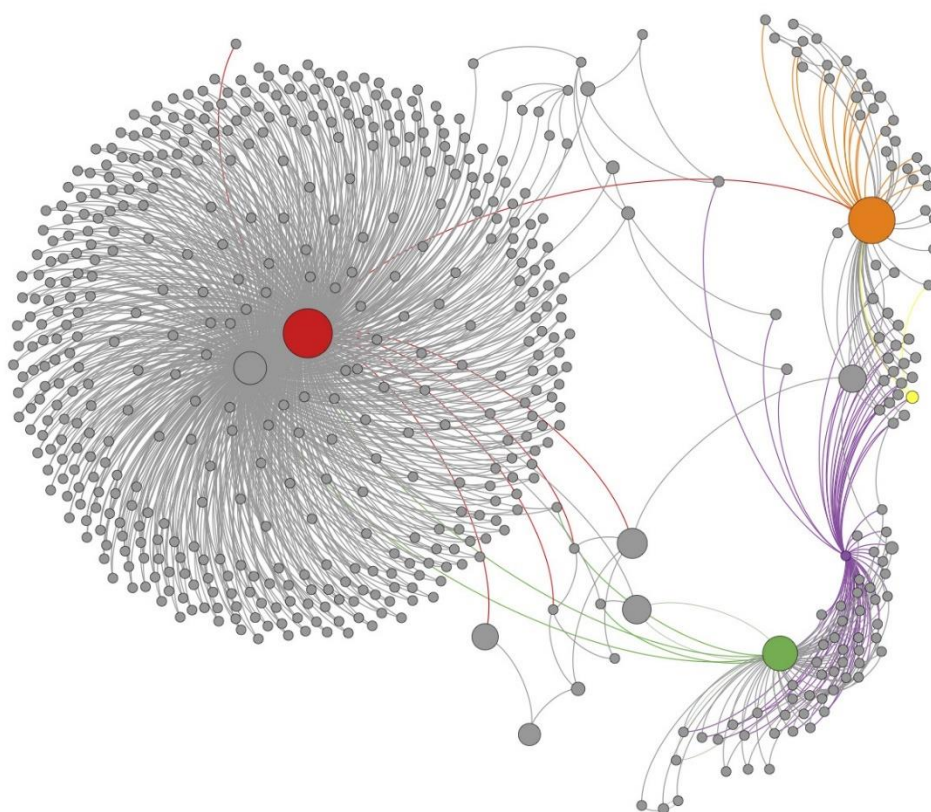
Δεδομένου ότι οι χρήστες μιας κοινότητας μοιράζονται κοινές ιδιότητες, η ανίχνευση κοινοτήτων σε ένα δίκτυο, αποτελεί ένα βασικό χαρακτηριστικό για την εξαγωγή χρήσιμων συμπερασμάτων της συνολικής λειτουργίας του δικτύου. Στην πραγματικότητα, ο τεράστιος όγκος δεδομένων απαιτεί μια αποτελεσματική μέθοδο για την ανίχνευση κοινοτήτων σε πολύπλοκα δίκτυα. Διάφορες μέθοδοι έχουν υλοποιηθεί με σκοπό την μέγιστη αποτελεσματικότητα, με την επικρατέστερη να είναι η αρθρωτότητα (modularity).

Η αρθρωτότητα αποτελεί μετρική αξιολόγησης για την διαμέριση του δικτύου σε κοινότητες σε σύγκριση με ένα τυχαίο δίκτυο, με τιμές από -1 έως 1. Όσο μεγαλύτερη η τιμή της αρθρωτότητας, τόσο καλύτερος ο διαχωρισμός του δικτύου.

Οι πρώτοι που υλοποίησαν την μέθοδο της αρθρωτότητας ήταν ο Girvan κ.α. [23] μέσω του αλγορίθμου Girvan-Newman. Ο αλγόριθμος Girvan-Newman, εφαρμόστηκε επιτυχώς σε πολλές περιπτώσεις για την ανίχνευση κοινοτήτων σε διάφορα δίκτυα παρέχοντας μια καλή λύση. Ωστόσο, υπάρχουν και άλλες μέθοδοι που δεν βασίζονται στην αρθρωτό-

τητα αλλά στην τυχειότητα, όπως ο αλγόριθμος Walktrap. Οι αλγόριθμοι αυτοί περιγράφονται στην ενότητα 5.1.

Για το δίκτυο Bitcoin, χρησιμοποιείται η παραδοσιακή μέθοδος ανίχνευσης μίας κοινότητας και εφαρμόζεται στο κοινωνικό δίκτυο συναλλαγών για την ομαδοποίηση χρηστών με παρόμοια χαρακτηριστικά. Έτσι, δημιουργείται ένα δίκτυο ως γράφημα όπου κάθε κόμβος είναι ένας χρήστης και για κάθε συναλλαγή προστίθενται μια ακμή από τον απόστολέα στον παραλήπτη. Παρόλου που το Bitcoin θεωρείται ανώνυμο, η ανίχνευση κοινοτήτων βοηθάει στην αναγνώριση των χρηστών και των συναλλαγών όπου αυτοί εμπειρεύονται. Για παράδειγμα, το 2011, εκλάπηκαν 25000 BTC και στάλθηκαν στην διεύθυνση 1KPTdMb6p7H3YCwsyFqrEmKGmsHqe1Q3jg [24]. Σύμφωνα με την ιχνηλάτηση των συναλλαγών, από την διεύθυνση αυτή, ένα ποσό στάλθηκε στην LulzSec (ομάδα χάκερ). Η εικόνα 40 παρουσιάζει τις συναλλαγές που πραγματοποιήθηκαν μετά την υποκλοπή. Από έρευνες στις κοινότητες αυτές, προέκυψε ότι υπήρξαν και άλλες περιπτώσεις υποκλοπής και μεταφοράς BTC σε public keys που ανήκουν στην ομάδα LulzSec.



Εικόνα 40: Δίκτυο συναλλαγών των χάκερ [25].

Κατά αυτόν τον τρόπο, υπάρχει η δυνατότητα ομαδοποίησης των public keys που ανήκουν σε αυτήν την ομάδα ή σε παρόμοιες ύποπτες συναλλαγές, με σκοπό την ανεπιτυχή πραγματοποίηση των συναλλαγών.

5.1 Αλγόριθμοι ανίχνευσης κοινοτήτων

Οι μέθοδοι ανίχνευσης κοινοτήτων κατηγοριοποιούνται σε:

- Ιεραρχική μέθοδος (agglomerative)
- Διαιρετική μέθοδος (divisive)

Στην ιεραρχική μέθοδο, οι ακμές μεταξύ των κόμβων προστίθενται μια προς μια από τον ένα κόμβο στον άλλο, ενώ στην περίπτωση της διαιρετικής υλοποίησης οι ακμές αφαιρούνται μια προς μια. Για την ανίχνευση κοινοτήτων σε ένα δίκτυο οι κύριοι αλγόριθμοι είναι:

- Girvan-Newman
- Louvain
- Leiden
- Walktrap

5.1.1 Girvan-Newman

Ο αλγόριθμος Girvan-Newman βασίζεται στην ιεραρχική διαιρετική μέθοδο αφαιρώντας ακμές από το δίκτυο σύμφωνα με την κεντρικότητα μεταξύ των άκμων (edge betweenness centrality) με βάση τα συντομότερα μονοπάτια (shortest paths). Για την ανίχνευση κοινοτήτων, ο αλγόριθμος υπολογίζει τα συντομότερα μονοπάτια σύμφωνα με τα κριτήρια συντομότερων μονοπατιών μεταξύ των κόμβων και στην συνέχεια υπολογίζει τις βαθμολογίες των ακμών. Η ακμή με την υψηλότερη βαθμολογία αφαιρείται από το δίκτυο και η διαδικασία επαναλαμβάνεται. Σε κάθε επανάληψη υπολογίζεται η αρθρωτότητα του δικτύου. Η μέγιστη αρθρωτότητα δίνει και την καλύτερη διαμέριση του δικτύου.

5.1.2 Louvain

Η υλοποίηση βασίζεται στην αρθρωτότητα του δικτύου, καθώς ο στόχος είναι η μεγιστοποίηση της διαφοράς μεταξύ των ακμών και των αναμενόμενων ακμών σε μια κοινότητα. Ο αλγόριθμος χωρίζεται σε δύο φάσεις. Στην πρώτη φάση, ο κάθε κόμβος αποτελεί μια κοινότητα. Στην συνέχεια, για κάθε κόμβο υπολογίζεται το κέρδος (gain) της αρθρωτότητας αφαιρώντας τον κόμβο αυτόν από την κοινότητα που βρίσκεται και προσθέτοντας τον στην κοινότητα του πιο κοντινού γείτονα. Στην περίπτωση που το κέρδος είναι θετικό και μεγιστοποιημένο, τότε ο κόμβος ομαδοποιείται με τον γείτονα του, ενώ

στην περίπτωση που το κέρδος είναι αρνητικό, παραμένει ως έχει. Η διαδικασία επαναλαμβάνεται για όλους τους κόμβους μέχρι την μεγιστοποίηση της αρθρωτότητας του τρέχων δικτύου. Στην δεύτερη φάση, ο αλγόριθμος δημιουργεί το δίκτυο με βάση τις κοινότητες που προέκυψαν από την πρώτη φάση. Τα βήματα αυτά επαναλαμβάνονται μέχρι να μην υπάρχει κάποια αλλαγή στο δίκτυο και η αρθρωτότητα να είναι η μέγιστη καλύτερη.

5.1.3 Leiden

Η υλοποίηση του αλγορίθμου Leiden είναι παρόμοια με την υλοποίηση του αλγορίθμου Louvain. Η διαφορά είναι στην δεύτερη φάση, όπου οι ήδη υπάρχουσες κοινότητες διαμερίζονται και κάποιοι κόμβοι ομαδοποιούνται τυχαία σε κοινότητες σύμφωνα με μια συνάρτηση ομαδοποίησης. Στην συνέχεια, η διαδικασία αυτή επαναλαμβάνεται μόνο για τους κόμβους που πραγματοποιήθηκε η ομαδοποίηση.

5.1.4 Walktrap

Ο αλγόριθμος Walktrap βασίζεται στην τυχαιότητα των μονοπατιών μεταξύ των κορυφών στο δίκτυο. Το κύριο χαρακτηριστικό του αλγορίθμου είναι ότι τα μονοπάτια κατευθύνονται τυχαία σε άμεσα συνδεδεμένους κόμβους, τις κοινότητες. Η διαμέριση του δικτύου αξιολογείται σύμφωνα με το κριτήριο της αρθρωτότητας του δικτύου.

Οι παραπάνω αλγόριθμοι εφαρμόζονται για την ανίχνευση κοινοτήτων στο δίκτυο Bitcoin. Πιο συγκεκριμένα, χρησιμοποιούνται για την καλύτερη διαμέριση του δικτύου με στόχο την εξαγωγή πληροφοριών και ομαδοποίησης των χρηστών σε διάφορες ομάδες με κοινά χαρακτηριστικά. Μετά την διαμέριση του δικτύου, είναι δυνατή η ανάλυση της λειτουργίας και συμπεριφοράς των χρηστών της κάθε ομάδας πιο εύκολα, καθώς άλλοι-λοεπιδρούν πιο συχνά μεταξύ τους. Ο τρόπος υλοποίησης περιγράφεται στο κεφάλαιο 6.

Κεφάλαιο 6. Πειραματική Αξιολόγηση

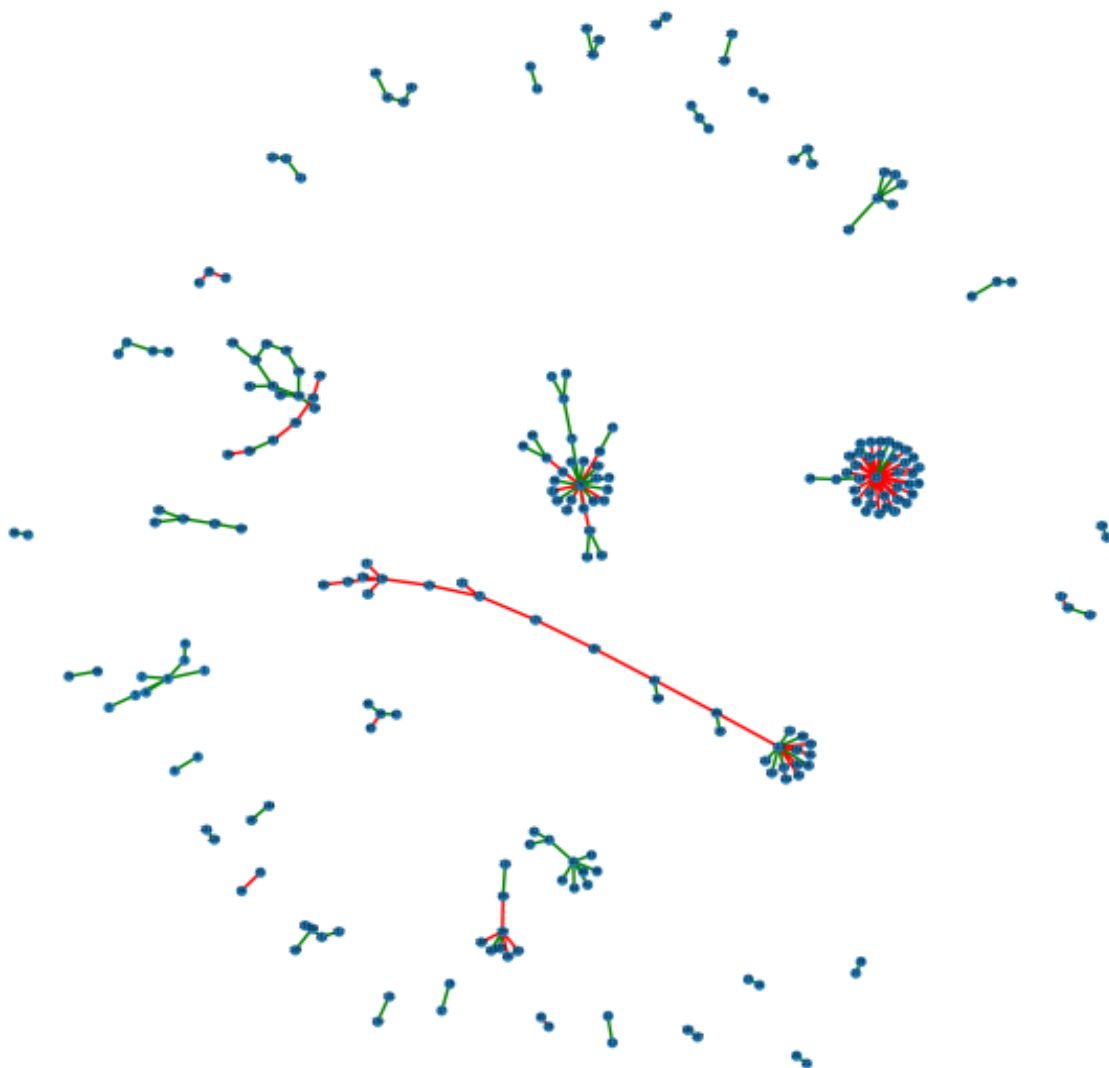
Σε αυτό το κεφάλαιο περιγράφεται η πειραματική αξιολόγηση των αλγορίθμων για την ανίχνευση κοινοτήτων. Μετά την ολοκλήρωση της αποκωδικοποίησης των δεδομένων και της ομαδοποίησης των διευθύνσεων, για κάθε συναλλαγή δημιουργείται το δίκτυο. Η αξιολόγηση πραγματοποιείται στις χρονολογίες 2009, 2010 όπου το Bitcoin βρίσκεται στο αρχικό στάδιο και την χρονολογία 2013 όπου παρατηρείται έξαρση. Συγκρίνεται η αποτελεσματικότητα κάποιων αλγορίθμων για την ανίχνευση κοινοτήτων και αναλύεται το συνολικό flow μεταξύ των χρηστών.

6.1 Προγραμματιστικά εργαλεία

Για την υλοποίηση των πειραμάτων χρησιμοποιήθηκε η γλώσσα προγραμματισμού Python καθώς και οι βιβλιοθήκες NetworkX, CDlib και Matplotlib. Η NetworkX χρησιμοποιήθηκε για την δημιουργία του γραφήματος, η Matplotlib για την σχεδίαση του γραφήματος, ενώ η βιβλιοθήκη CDlib περιέχει τους αλγόριθμους για την ανίχνευση κοινοτήτων.

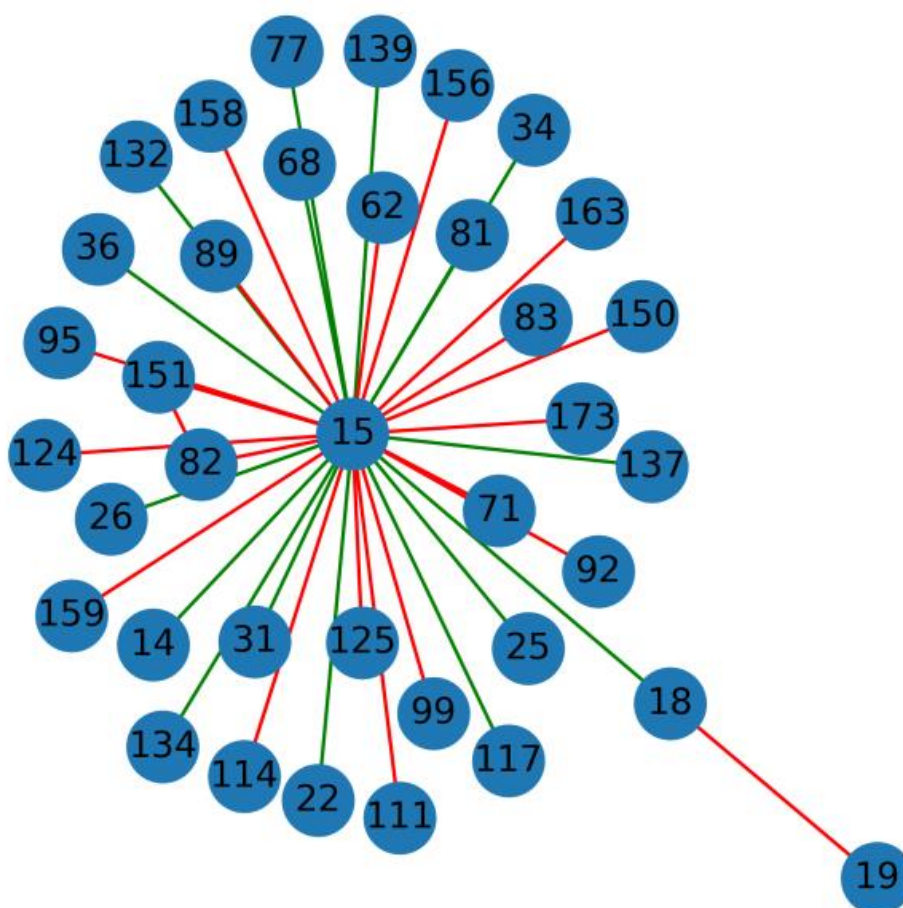
6.2 Ανάλυση αποτελεσμάτων

Στην παρακάτω εικόνα απεικονίζονται οι κοινότητες για το έτος 2009. Το δίκτυο αποτελείται από 40 κοινότητες, 222 χρήστες και 208 αλληλεπιδράσεις. Οι κόκκινες ακμές δηλώνουν συναλλαγές που πραγματοποιήθηκαν με μεταφορά αξίας μεγαλύτερη από 500 BTC ενώ αντίστοιχα οι υπόλοιπες συναλλαγές απεικονίζονται με πράσινες ακμές.



Εικόνα 41: Απεικόνιση κοινοτήτων για το έτος 2009.

Για παράδειγμα, στην παρακάτω κοινότητα παρατηρούμε ότι πραγματοποιήθηκαν 21 συναλλαγές μεταξύ των χρηστών μεταφέροντας ποσό αξίας μεγαλύτερο από 500 BTC.



Εικόνα 42: Παράδειγμα αλληλοεπιδράσεων μιας κοινότητας για το έτος 2009.

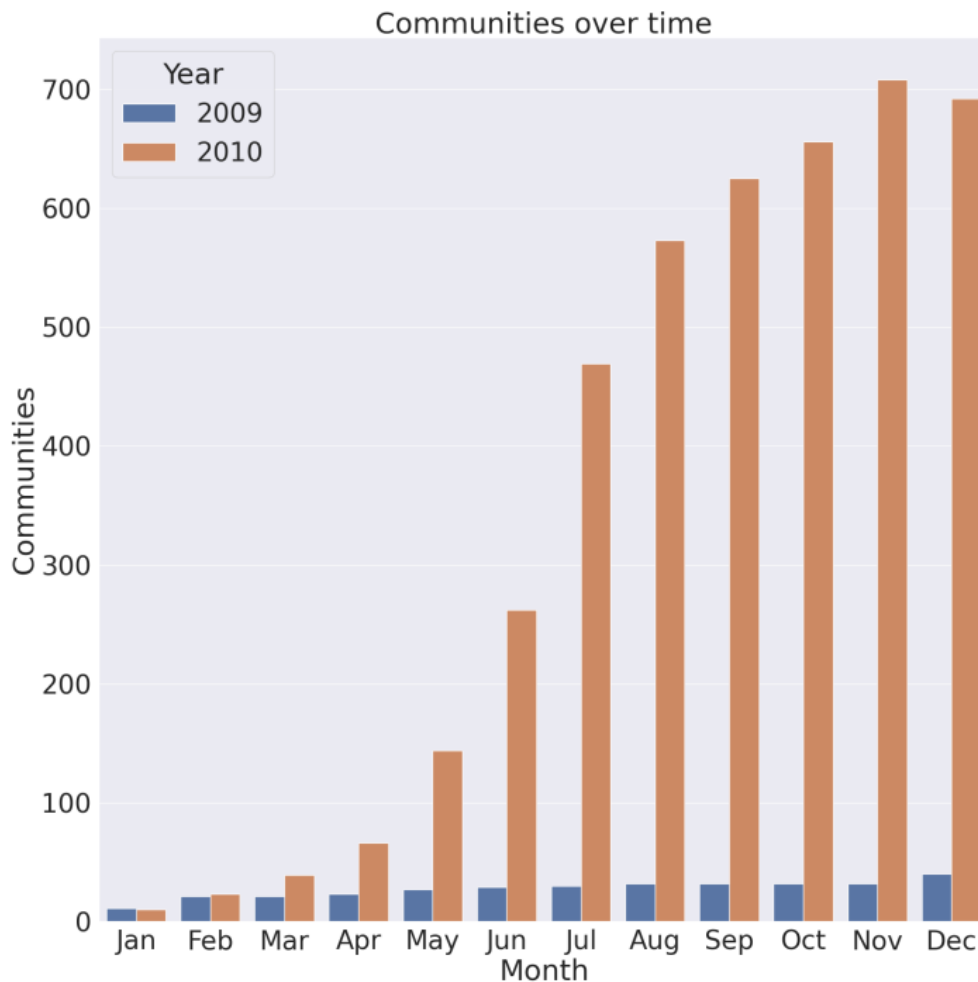
Για την αποφυγή του bottleneck, η σύγκριση της αποτελεσματικότητας των αλγορίθμων πραγματοποιείται στις χρονολογίες 2009 και 2010, καθώς για το έτος 2013 το δίκτυο είναι αρκετά μεγάλο.

Πίνακας 4: Ανάλυση δικτύου για χρονολογίες 2009, 2010 και 2013.

Έτος	2009	2010	2013
Χρήστες	222	46.016	4.662.681
Αλληλοεπιδράσεις	208	98.868	26.727.323
Συνολικό flow (BTC)	277.419	16.056.107	219.302.499
Κοινότητες	40	692	-

Στην παρακάτω εικόνα απεικονίζονται οι κοινότητες και πως αλλάζουν σε βάθος χρόνου. Τον πρώτο χρόνο (2009) λειτουργίας του Bitcoin, δεν παρατηρείται μεγάλη αύξηση στο σύνολο των κοινοτήτων αλλά ούτε και στον αριθμό των αλληλοεπιδράσεων μεταξύ των

χρηστών καθώς βρίσκεται στο αρχικό στάδιο και δεν είναι ακόμη ευρέως διαδεδομένο. Κάτι το οποίο αλλάζει το 2010, και πραγματοποιούνται περισσότερες συναλλαγές. Κατά συνέπεια, αυξάνει ο αριθμός των χρηστών και των αλληλοεπιδράσεων με αποτέλεσμα να αυξάνει ο αριθμός των κοινοτήτων.



Εικόνα 43: Σύγκριση κοινοτήτων για το 2009 και 2010.

Στους πίνακες 5 και 6 γίνεται σύγκριση της αποτελεσματικότητας των αλγορίθμων στην συνολική διαμέριση του δικτύου τον μήνα Δεκέμβριο για τις χρονολογίες 2009 και 2010 αντίστοιχα. Οι κοινότητες για το έτος 2009 είναι 40 ενώ για το έτος 2010 είναι 692. Για το έτος 2009, επειδή ο αριθμός των κοινοτήτων είναι μικρός, όλοι οι αλγόριθμοι υπολογίζουν την ίδια διαμέριση του δικτύου, προσφέροντας μια καλή λύση. Σε αντίθεση για το έτος 2010, οι αλγόριθμοι Lounvain και Leiden φαίνεται να είναι πιο αποτελεσματικοί (σε δίκτυα μεγάλης κλίμακας) σε σχέση με τον αλγόριθμο τυχειότητας Walktrap, με βάση

την τιμή της αρθρωτότητας (όσο πιο κοντά στο 1, τόσο καλύτερη διαμέριση του δικτύου). Ενώ ο Girvan-Newman επειδή υπολογίζει το edge betweenness για κάθε ακμή αποφεύγεται λόγω υψηλής υπολογιστικής πολυπλοκότητας καθώς τρέχει σε $O(n*m)$.

Πίνακας 5: Σύγκριση αποτελεσματικότητας αλγορίθμων για το 2009

Αλγόριθμος	Διαμέριση του δικτύου σε κοινότητες	Αρθρωτότητα	Χρόνος (sec)
Girvan-Newman	42	0.909	0.039
Louvain	42	0.909	0.184
Leiden	42	0.909	0.007
Walktrap	42	0.909	0.006

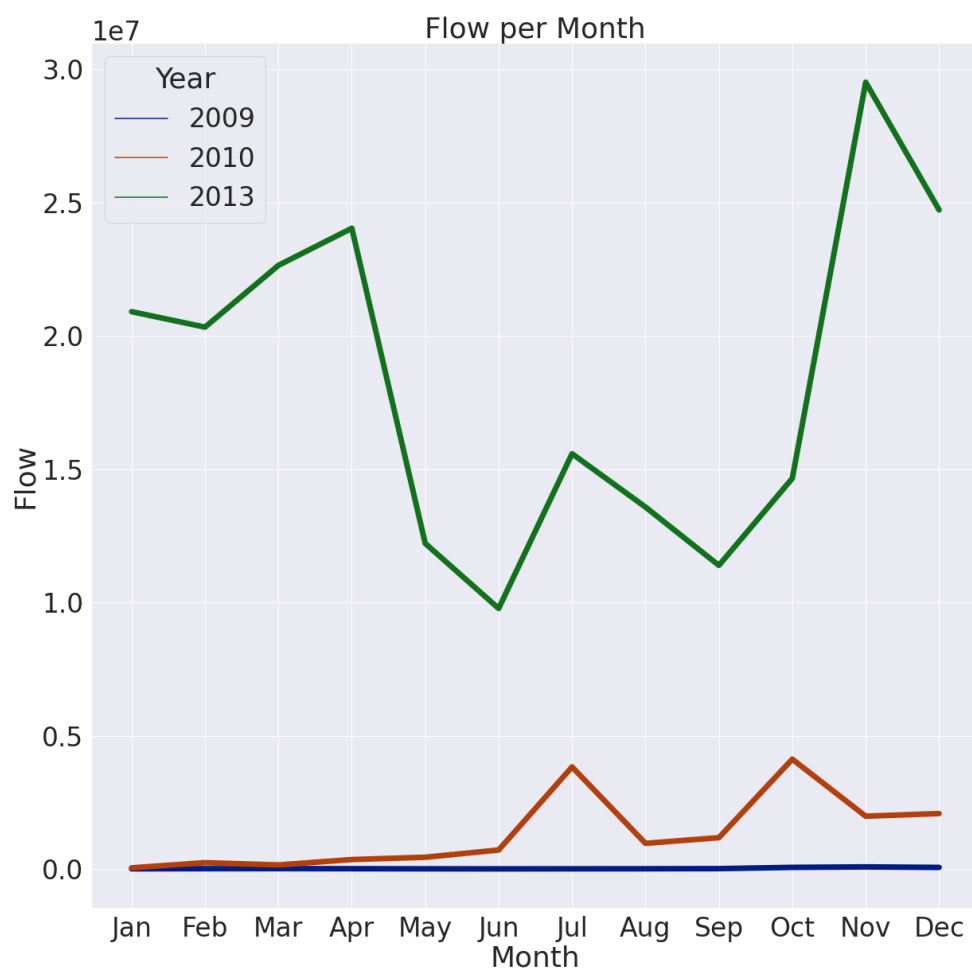
Πίνακας 6: Σύγκριση αποτελεσματικότητας αλγορίθμων για το 2010

Αλγόριθμος	Διαμέριση του δικτύου σε κοινότητες	Αρθρωτότητα	Χρόνος (sec)
Louvain	783	0.850	28.57
Leiden	776	0.858	1.23
Walktrap	5702	0.685	51.38

Ο πίνακας 6 περιέχει πληροφορίες για την αξία BTC που στάλθηκε. Παρατηρείται ότι οι συναλλαγές αυξάνονται με γεωμετρική πρόοδο, ειδικά το έτος 2013 όπου το Bitcoin βρίσκεται σε έξαρση. Το 2009, 75 χρήστες έλαβαν BTC αξίας μικρότερης από 50, ενώ το 2013 οι χρήστες ήταν 26.148.392(!). Η διαφορά είναι εμφανής και παρατηρείται επίσης στην εικόνα 48 όπου απεικονίζεται το συνολικό flow ανά μήνα και έτος.

Πίνακας 6: Απεικόνιση αξίας BTC που έλαβαν οι χρήστες.

Αξία σε BTC	2009	2010	2013
<50	75	80.665	26.148.392
50-150	22	5.871	353.498
150-250	12	1.862	79.826
250-500	28	4.850	77.108
>500	71	5.620	68.499



Εικόνα 44: Σύγκριση flow ανά μήνα (όπου 0.5 = 5.000.000).

Κεφάλαιο 7. Επίλογος

Ο στόχος της διπλωματικής εργασίας ήταν η αποκωδικοποίηση των συναλλαγών του δικτύου Bitcoin, η ανάλυση των χρηστών και των αλληλοεπιδράσεων μεταξύ τους καθώς και η ανίχνευση των κοινοτήτων που απαρτίζουν το δίκτυο. Κατά την αποκωδικοποίηση, χρησιμοποιήθηκαν βάσεις δεδομένων για καλύτερη και πιο αποτελεσματική επεξεργασία των δεδομένων, ενώ στην συνέχεια δημιουργήθηκε ένα δίκτυο που αποτελείται από τις συναλλαγές μεταξύ των χρηστών. Κατά την ανάλυση του δικτύου και σύμφωνα με την αξιολόγηση των πειραμάτων, είναι εμφανές ότι το Bitcoin έχει μπει για τα καλά στην οικονομία ως ένα ευρέως χρησιμοποιούμενο εναλλακτικό ιδιωτικό νόμισμα. Με την πάροδο των χρόνων, χρησιμοποιείται ακόμα πιο πολύ για διάφορες πληρωμές ή για την αγορά φυσικών αγαθών και υπηρεσιών. Κατά συνέπεια, ο αριθμός των αλληλοεπιδράσεων μεταξύ των χρηστών αυξάνει ραγδαία με αποτέλεσμα τον σχηματισμό περισσότερων κοινοτήτων που συνθέτουν το δίκτυο.

Μελλοντικά, θα μπορούσαμε να κατευθυνθούμε στα εξής:

- Καθώς παρατηρείται αύξηση στον αριθμό των κοινοτήτων, αυξάνει και η πολυπλοκότητα του δικτύου, με αποτέλεσμα την ανάγκη μιας μεθόδου υλοποίησης που θα παρέχει μια καλή διαμέριση του δικτύου, καθώς και μείωση του κόστους της υπολογιστικής πολυπλοκότητας.
- Σε ενδεχόμενη αναβάθμιση του πρωτοκόλλου σχετικά με τα blocks, όπως στην περίπτωση του δικτύου Lightning, ο κώδικας θα πρέπει να τροποποιηθεί κατάλληλα για την σωστή αποκωδικοποίηση των δεδομένων.
- Καθώς οι συναλλαγές είναι δημόσιες, θα μπορούσε να υλοποιηθεί μια μέθοδος web scraping για την αποθήκευση των συναλλαγών από μια online πηγή, αντί για την εγκατάσταση του λογισμικού Bitcoin Core.

Βιβλιογραφία

Papers

- [MaZh17] D. Mao, Y. Zhang. User Categorization and Community Detection in Bitcoin Network, pp. 2017.
- [Naka08] S. Nakamoto. Bitcoin: A Peer-to-Peer Electronic Cash System, pp. 2008.
- [ReRM10] C. Remy, B. Rym, L. Matthieu Tracking bitcoin users activity using community detection on a network of weak signals, pp. 2017.
- [YuBu15] A. Yu, B. Bunz. Community Detection and Analysis in the Bitcoin Network, CS 224W Final Report, pp. 2015.

Web articles & images

- [1] [Peer-to-peer - Wikipedia](#)
- [2] [Proof of work - Wikipedia](#)
- [3] [Blockchain - Wikipedia](#)
- [4] [The Untold Story of Silk Road, Part 1 | WIRED](#)
- [5] [blockchain images - Αναζήτηση Google](#)
- [6] [Blockchain - Hash of previous block! — Steemit](#)
- [7] [Blockchain for Dummies. The five keys to understanding what is... | by Telmo Subira Rodriguez | The Startup | Medium](#)
- [8] [What is Proof of Work \(PoW\) | Definition and Meaning | Capital.com](#)
- [9] [Elliptic Curve Digital Signature Algorithm - Wikipedia](#)
- [10] [What Are Public Keys and Private Keys? | Ledger](#)
- [11] [File:Public key encryption alice to bob.svg - Wikimedia Commons](#)
- [12] [en-transaction-propagation.png \(800×627\) \(wp.com\)](#)
- [13] [The Best Step-by-Step Bitcoin Script Guide: Part One \(Blockgeeks\)](#)
- [14] [Ch.10: Something on Transaction Unlocking & Locking Script | by Khor Aik Cheow, PhD | Medium](#)
- [15] [P2WPKH \(Pay to Witness Public Key Hash\) - Programming The Blockchain in C# \(gitbook.io\)](#)

- [16] [Bitcoin to Rand – Bitcoin market Guide South Africa \(bitcoin-guide-africa.com\)](http://bitcoin-guide-africa.com)
- [17] [Blockchain Explorer - Search the Blockchain | BTC | ETH | BCH](#)
- [18] [Master's Thesis : BitCoin Clustering: a CoinJoin Discarding Heuristic \(uliege.be\)](#)
- [19] [What is a Block in BlockChain? | SAP Blogs](#)
- [20] [Transaction Receipt](#)
- [21] [Transaction:
f4184fc596403b9d638783cf57adfe4c75c605f6356fbc91338530e9831e9e16 | Blockchain Explorer](#)
- [22] [Community Detection Algorithms. Many of you are familiar with networks... | by Thamindu Dilshan Jayawickrama | Towards Data Science](#)
- [23] [Girvan–Newman algorithm - Wikipedia](#)
- [24] [Address: 1KPTdMb6p7H3YCwsyFqrEmKGmsHqe1Q3jg](#)
- [25] [An Analysis of Anonymity in the Bitcoin System: Bitcoin is not Anonymous \(anonymity-in-bitcoin.blogspot.com\)](#)