Θεωρία Παιγνίων και Κίνητρα στο Blockchain: Δικαιοσύνη, Προστριβές και Συμμετοχή

Γιώργος Τσούμας

Universitat Pompeu Fabra Barcelona

OdyC

Περιεχόμενα

1

Εισαγωγή

Blockchain
Σχεδιασμός Μηχανισμών
(Mechanism Design)

Bitcoin: A Peer-to-Peer Electronic Cash System

Satoshi Nakamoto
satoshin@gmx.com
www.bitcoin.org

Abstract. A purely peer-to-peer version of electronic cash would allow online payments to be sent directly from one party to another without going through a financial institution. Digital signatures provide part of the solution, but the main benefits are lost if a trusted third party is still required to prevent double-spending.

Το 2008, ένα άτομο με το ψευδώνυμο **Satoshi Nakamoto** δημοσίευσε το whitepaper του **Bitcoin** [1]. Στόχος του είναι να αντικατασταθεί το παραδοσιακό τραπεζικό σύστημα με ένα **αποκεντρωμένο δίκτυο ηλεκτρονικών πληρωμών**.

Ορισμοί

Αποκεντροποιημένο λογιστικό βιβλίο, η τήρηση του οποίου γίνεται με την βοήθεια του blockchain.

Μια **δομή δεδομένων**, η οποία διατηρείται από τους κόμβους ενός κατανεμημένου δικτύου.

Το σύνολο των **στοιχείων** που διατηρούνται πάνω στο Ledger.

Προγράμματα που εκτελούνται μέσω του blockchain. Πολλές φορές παίζουν τον ρόλο του **διαμεσολαβητή**.

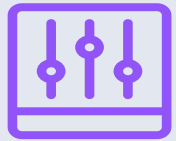
Ανατομία του Blockchain

Δίκτυο

Οι κόμβοι (nodes) επικοινωνούν μεταξύ τους για να διατηρούν το κοινό λογιστικό βιβλίο (ledger) και να επαληθεύουν τις συναλλαγές.

Ομοφονία

Είναι η διαδικασία με την οποία όλοι οι κόμβοι του δικτύου συμφωνούν για την εγκυρότητα των συναλλαγών



dApps

Εφαρμογές που λειτουργούν πάνω σε blockchain, χρησιμοποιώντας έξυπνα συμβόλαια για την εκτέλεση λειτουργιών

Παίγνια και ισορροπία Nash

Ένα **παιγνίο** είναι ένα μοντέλο αλληλεπίδρασης μεταξύ περισσοτέρων του ενός **λογικών** και **στρατηγικά σκεπτόμενων** παικτών, όπου η απόφαση του κάθε παίκτη επηρεάζει το τελικό αποτέλεσμα.

Ένα παιγνίο αποτελείται από:

- Ένα σύνολο παικτών N={1,2,..n}.
- Ένα σύνολο στρατηγικών για κάθε παίκτη
- Μία συνάρτηση πληρωμής (**utility**) για κάθε παίκτη, η οποία καθορίζει την απόδοσή του ανάλογα με τις στρατηγικές όλων των παικτών.

Ισορροπία Nash είναι ένα προφίλ στρατηγικών όπου κανένας παίκτης δεν μπορεί να βελτιώσει την πληρωμή του αλλάζοντας μονομερώς τη στρατηγική του, δεδομένων των στρατηγικών των άλλων παικτών. Μαθηματικά:

$$u_i(s_i^*, s_{-i}^*) \geq u_i(s_i, s_{-i}^*) \quad \forall \quad s_i \in S_i$$

Σχεδιασμός Μηχανισμών (Mechanism Design). Η θεωρία παιγνίων από την ανάποδη. Σχεδιασμός μηχανισμών για να εξασφαλίσουμε καλύτερα κοινωνικά αποτελέσματα ενώ οι συμμετέχοντες έχουν ιδιωτικά συμφέροντα.

Ένας μηχανισμός περιλαμβάνει:

- **Έναν χώρο στρατηγικών** για κάθε παίκτη
- **Μία συνάρτηση διαμοιρασμού** που αντιστοιχίζει τις στρατηγικές σε ένα τελικό αποτέλεσμα (π.χ. κατανομή αγαθών, τιμές, αποφάσεις),
- Συχνά και **κανόνες πληρωμής ή ανταμοιβής**, ειδικά σε οικονομικά περιβάλλοντα.

Αποτέλεσμα για κάθε παίκτη: $u_i\big(x(s_1,\ldots,s_n)\big) = v_i\big(x(s_1,\ldots,s_n)\big) - p_i(s_1,\ldots,s_n)$

Σχεδιασμός Μηχανισμών (Mechanism Design). Η θεωρία παιγνίων από την ανάποδη. Σχεδιασμός μηχανισμών για να εξασφαλίσουμε καλύτερα κοινωνικά αποτελέσματα ενώ οι συμμετέχοντες έχουν ιδιωτικά συμφέροντα.

Ένας μηχανισμός περιλαμβάνει:

- **Έναν χώρο στρατηγικών** για κάθε παίκτη
- **Μία συνάρτηση διαμοιρασμού** που αντιστοιχίζει τις στρατηγικές σε ένα τελικό αποτέλεσμα (π.χ. κατανομή αγαθών, τιμές, αποφάσεις),
- Συχνά και **κανόνες πληρωμής ή ανταμοιβής**, ειδικά σε οικονομικά περιβάλλοντα.

Αποτέλεσμα για κάθε παίκτη: $u_i\big(x(s_1,\ldots,s_n)\big) = v_i\big(x(s_1,\ldots,s_n)\big) - p_i(s_1,\ldots,s_n)$

Ιδιότητες ενός "Καταπληκτικού" Μηχανισμού (Awesome Mechanism Design – Roughgarden*)

- **Συμβατότητα Κινήτρων** (Incentive Compatibility)
- **Μέγιστη Κοινωνική Ευημερία** (Social Welfare Maximization)
- **Αποδοτικότητα** (Efficiency)
- **Πρακτικότητα** (Practicality)
- **Δικαιοσύνη** (Fairness)
- **Ατομική Ορθολογικότητα** (Individual Rationality)
- **Ανθεκτικότητα** (Robustness)

* Gödel Prize 2012

2

Flash Boys 2.0*

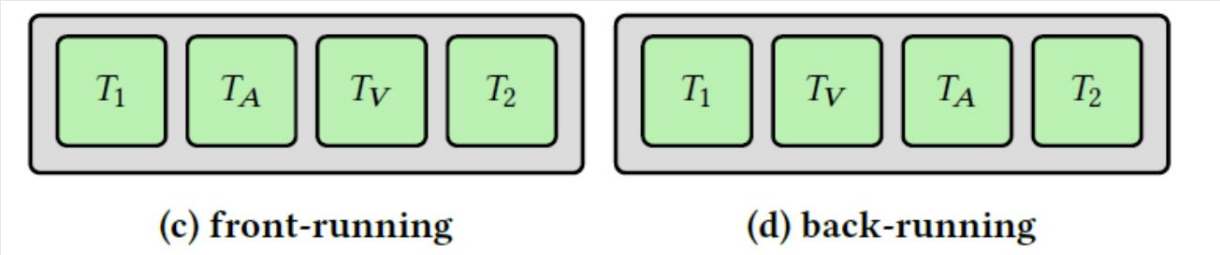
ethereum

Ο όρος **"Flash Boys"** στο πλαίσιο των χρηματιστηρίων αναφέρονται σε **επενδυτές υψηλής συχνότητας (High-Frequency Traders - HFTs)** που χρησιμοποιούν **υπερταχείς αλγορίθμους και υποδομές** για να εκμεταλλευτούν μικρές ανισορροπίες στις χρηματοπιστωτικές αγορές — συχνά εις βάρος των απλών επενδυτών.

Στο **Ethereum**[2], κάθε νέα σελίδα του ledger (block) παράγεται ως εξής: οι χρήστες στέλνουν τις συναλλαγές τους, και έπειτα ένας κόμβος διαλέγει ποιες από αυτές θα συμπεριληφθούν στο block **και με ποιά σειρά**, με βάση τα tips που αφήνουν οι χρήστες στον κόμβο. Εάν η ζήτηση εκείνη τη στιγμή είναι υψηλή, κάποιες από αυτές ίσως μείνουν εκτός.

Στο **Ethereum**[2], κάθε νέα σελίδα του ledger (block) παράγεται ως εξής: οι χρήστες στέλνουν τις συναλλαγές τους, και έπειτα ένας κόμβος διαλέγει ποιες από αυτές θα συμπεριληφθούν στο block **και με ποιά σειρά**, με βάση τα tips που αφήνουν οι χρήστες στον κόμβο. Εάν η ζήτηση εκείνη τη στιγμή είναι υψηλή, κάποιες από αυτές ίσως μείνουν εκτός.

Το Front-running (αντίστοιχα Back-running) είναι μια αναδιάταξη της σειράς των προς εκτέλεση συναλλαγών, στην οποία η συναλλαγή του επιτιθέμενου εκτελείται πριν (αντίστοιχα μετά) την συναλλαγή του θύματος. Αφορά κυρίως συναλλαγές DEX*.

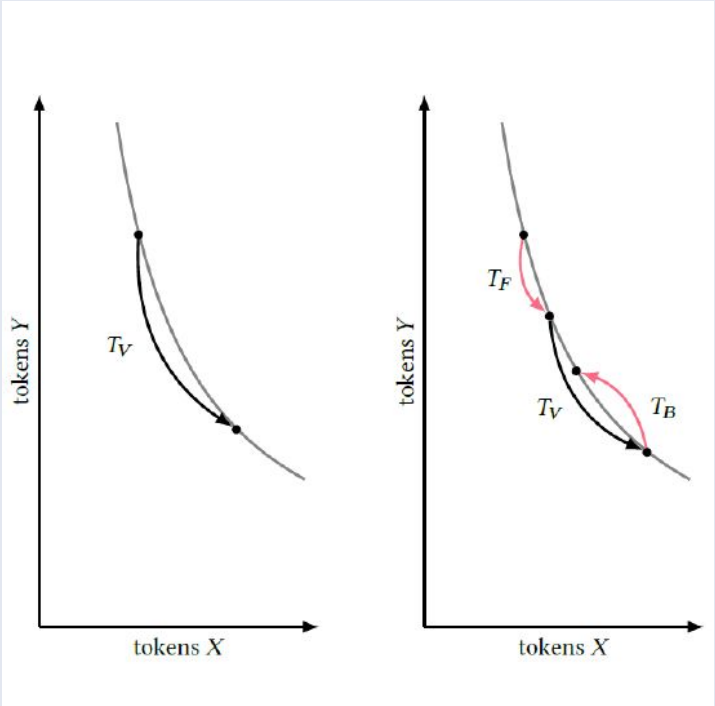


*Τα DEX (Decentralized Exchanges) είναι αποκεντρωμένα ανταλλακτήρια assets, πίσω από τα οποία δεν υπάρχει κάποια κεντρική αρχή που τα ελέγχει. Οι ανταλλαγές και η παροχή ρευστότητας γίνονται από τους ίδιους τους χρήστες, με τρόπο διαφανή και απόλυτα καθορισμένο.

Οι επιθέσεις "sandwich" συνδυάζουν τις συναλλαγές front-running και back-running για να "σαντουιτσάρουν" τη συναλλαγή του θύματος. Ο επιτηθέμενος εκτελεί την εξής στρατηγική:

- Δημιουργεί μια εντολή αγοράς, για να φουσκώσει την τιμή του asset Y.
- Συμπεριλαμβάνει την εντολή αγοράς του trader, φουσκώνοντας περαιτέρω την τιμή του asset Y.
- Τέλος, συμπεριλαμβάνει την εντολή πώλησης τους, που πουλάει όλες τις μονάδες του asset Y.

Το παραπάνω αποτελεί μια ευκαιρία καθαρού κέρδους (**Pure Revenue Opportunity**).



Οι Daian et al. στο [3] μοντελοποίησαν το προηγούμενο σενάριο στο παρακάτω παίγνιο, με βάση τις συμπεριφορές που είχαν παρατηρηθεί, καθώς και τους τεχνικούς κανόνες λειτουργίας του Ethereum.

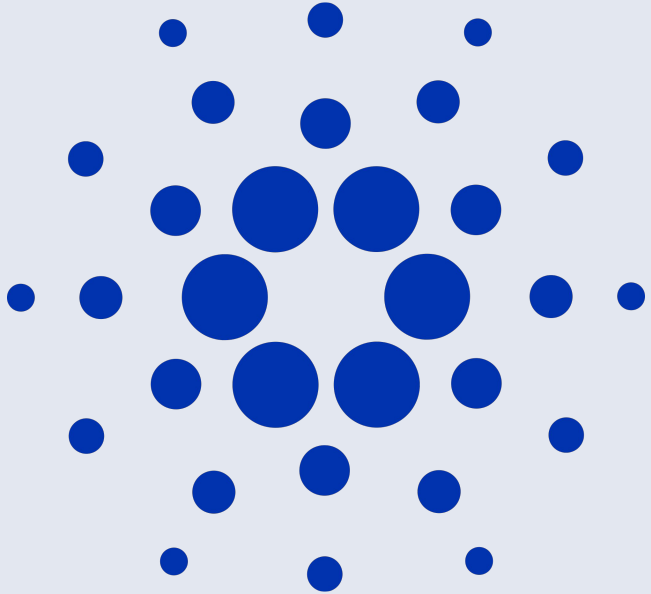
- **Συνεχής Χρόνος**.
- **Ατελής Πληροφόρηση**: Οι παίκτες παρατηρούν τις προσφορές με καθυστέρηση.
- **Όλοι Πληρώνουν**: Οι ηττημένοι παίκτες πληρώνουν τα κόστη για αποτυχημένες συναλλαγές.
- **Πιθανότητα Διάρκειας Δημοπρασίας**: Οι δημοπρασίες τερματίζονται τυχαία όταν εξορυχτεί το επόμενο μπλοκ.
- **Περιορισμένη Υποβολή Προσφορών**: Οι παίκτες πρέπει να περιμένουν ένα μικρό διάστημα πριν αυξήσουν τις προσφορές τους.
- **Ελάχιστη Αρχική Προσφορά**.
- **Ελάχιστες Αυξήσεις Προσφορών**: Οι παίκτες μπορούν να αυξήσουν τις προσφορές τους μόνο με ελάχιστο όριο.

Οι Daian et al. απέδειξαν πως για το παραπάνω παίγνιο, στη συνεργατική του μορφή, υπάρχει μια grim trigger ισορροπία Nash.

- **Συνεργατική**: Οι παίκτες υποβάλλουν bids με τη σειρά τους, ώστε τα συνολικά κόστη για όλους τους παίκτες να παραμείνουν χαμηλά. Επίσης, ανεβάζουν τα bids τους σταδιακά.
- **Grim trigger**: Όταν ένας παίκτης σπάσει την παραπάνω σύμβαση (π.χ. bidάροντας πιο ψηλά από ό,τι του αναλογεί), ο άλλος "τραβάει τη σκανδάλη" και bidάρει όσο τον value του PRO, ουσιαστικά "σκοτώνοντας" την ευκαιρία και για τους δύο.

3

Reward Sharing Schemes for Stake Pools



Ο Girolamo Cardano ήταν Ιταλός μαθηματικός, γιατρός, φιλοσόφος και μηχανικός, ο οποίος γεννήθηκε το 1501 και πέθανε το 1576. Είναι γνωστός για τη σημαντική του συνεισφορά στα μαθηματικά, τη θεωρία των πιθανοτήτων, και τη μηχανική.

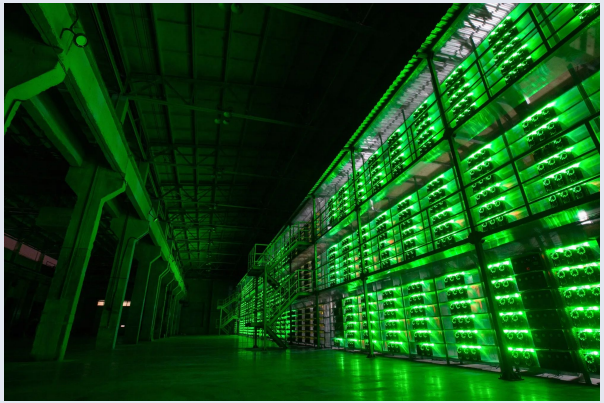
Το incentive structure του Bitcoin έχει ένα μηριαίο πρόβλημα. Ο τρόπος διαμοιρασμού των rewards στους miners αυξάνεται γραμμικά με το ποσοστό των συνολικών resources (υπολογιστική δύναμη) που ο καθένας τους ελέγχει. Άρα οι miners έχουν ισχυρά κίνητρα να δημιουργήσουν όλο και ισχυρότερα coalitions (mining pools). Αποτέλεσμα: κεντροποίηση.

Το Cardano[4] έλυσε αυτό το πρόβλημα χρησιμοποιώντας το **Reward Sharing Schemes for Stake Pools**[5] των Brünjes, Κιαγία, Κουτσουπιά* και Στούκα.

* Gödel Prize 2012

Τι είναι μια Επίθεση Sybil;

- **Ορισμός**: Μια επίθεση Sybil συμβαίνει όταν ένας μόνο αντίπαλος δημιουργεί πολλαπλές ψεύτικες ταυτότητες σε ένα δίκτυο για να αποκτήσει δυσανάλογη επιρροή.
- **Επίδραση**: Αυτό μπορεί να διαταράξει τους μηχανισμούς συναίνεσης, να χειραγωγήσει δεδομένα ή να εκκινήσει επιθέσεις στο δίκτυο.
- **Απάντηση**: Proof of Work, Proof of Stake



Κάθε συμμετέχοντας παίκτης (stakeholder) έχει από ένα μερίδιο του συστήματος (stake). Για την ορθή εκτέλεση του πρωτοκόλου απαιτούνται τα ακόλουθα:

- **Αποκεντροποίηση**: το ποσοστό των μεριδίων των τίμιων παικτών να είναι μεγαλύτερο του 51%, το οποίο δε, πρέπει να είναι κατανεμημένο σε έναν αρκετά μεγάλο αριθμό διαφορετικών παικτών.

- **Sybil Resistance**: οι κακόβουλοι παίκτες δεν θα πρέπει να μπορούν να πάρουν πλεονέκτιμα έναντι των τίμιων από την δημιουργία ”ψεύτικων προφίλ”.

Κάθε παίκτης μπορεί να συμμετάσχει ο ίδιος ή να μεταβιβάσει το μερίδιό του σε έναν άλλον ενεργό παίκτη. ΄Ετσι δημιουργούνται **συνασπισμοί** παικτών (pools), που στον καθέναν από αυτούς ηγείτε ένας ενεργός παίκτης.

Στόχος: Η ανάπτυξη ενός μηχανισμού διαμοιρασμού ενός ποσού R, ο οποίος δίνει κίνητρο στους παίκτες να εκτελέσουν το πρωτόκολο **τίμια** και **αποκεντροποιημένα**.

Το παίγνιο

- n παίκτες
- Ένα διάνυσμα μεριδίων, $s = (s_1, \ldots, s_n) : \sum_{i=1}^{n} = 1$
- Ένα διάνυσμα στατηγικών, $\forall$ παίκτη i
  $\vec{a_i} = (a_{i,1} \cdots, a_{i,n}) : \sum_{j}^{n} a_{i,j} = s_i$
- Ένα ιδιωτικό διάνυσμα κοστών $c = c_1, \ldots, c_n$, για όποιους παίκτες διαλέξουν να συμμετάσχουν ενεργά
- Για κάθε ενεργό παίκτη i, σ_i το συνολικό stake του συνασπισμού που ηγείται.
- Η παράμετρος $k \in \mathbf{N}, k < n$, το επιθυμητό πλήθος των ενεργών παικτών, και $\beta = \frac{1}{k}$
- Η Sybil παράμετρος $\alpha \in [0, \infty)$

Πρώτη προσπάθεια

Για του ηγέτες i κάθε συνασπισμού:

$$u_{i,i} = \begin{cases} r(\sigma_i, a_{i,i}) - c_i & \text{εάν } r(\sigma_i, a_{i,i}) \leq c_i \\ \frac{a_{i,i}}{\sigma_i} \cdot (r(\sigma_i, a_{i,i}) - c_i) & \text{διαφορετικά} \end{cases}$$

Ενώ για τους υπόλοιπους j $\neq$i:

$$u_{i,i} = \begin{cases} 0 & \text{εάν } r(\sigma_i, a_{i,i} \leq c_i \\ \frac{a_{j,i}}{\sigma_i} \cdot (r(\sigma_i, a_{i,i}) - c_i) & \text{διαφορετικά} \end{cases}$$

Όπου $\sum_{i=1}^{n} r(\sigma_i, a_{i,i}) \leq R$ και $r(0,0) = 0$.

Για τις παραπάνω (γνησίως αύξουσες) συναρτήσεις ωφέλειας, δεν υπάρχει σημείο ισορροπίας για το οποίο υπάρχουν πάνω από ένας συνασπισμοί.

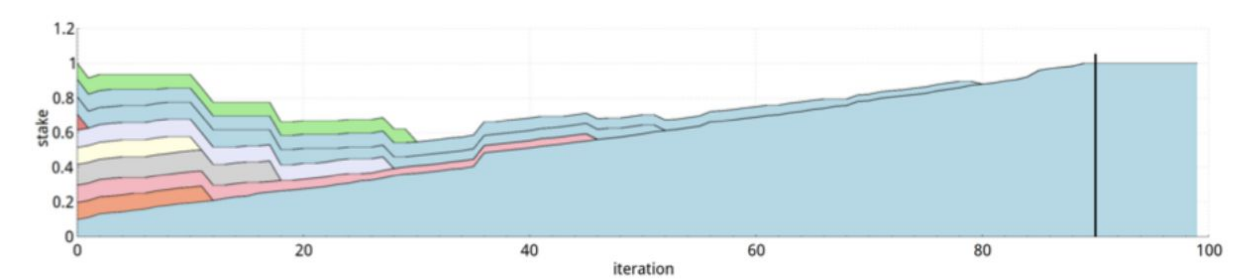
- **Cap**: Η συνάρτηση ανταμοιβής είναι αύξουσα μέχρι ενός σημείου συνολικού stake $\forall\lambda \quad r(\sigma,\lambda) = r(\beta,\lambda), \quad \text{for } \sigma > \beta$
- **Margin**: Η ηγέτης λαμβάνει ένα έξτρα μερίδιο των κερδών του συνασπισμού
- $$\frac{d(r(\sigma,\lambda) - c) \cdot \left(\frac{1}{\sigma}\right)}{d\sigma} > 0, \quad \sigma \leq \beta$$

Η συνάρτηση ανταμοιβής αυξάνεται για μικρές τιμές του stake του pool, ώστε να δοθούν κίνητρα στους παίκτες να συνεργαστούν και να μοιραστούν το κόστος

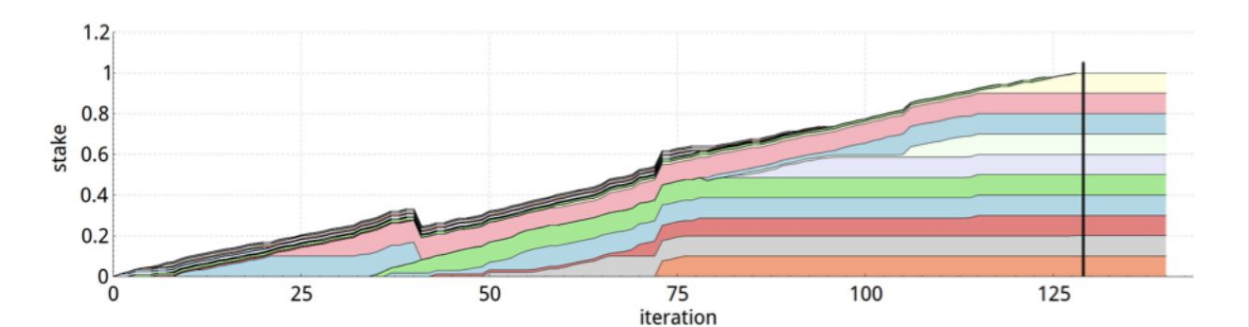
Το τελικό σχήμα

$$r(\sigma, \lambda) = \frac{R}{1+\alpha} \cdot (\sigma' + \lambda' \cdot \alpha \cdot \frac{\sigma' - \lambda' \cdot (1 - \sigma'/\beta)}{\beta})$$

Όπου $\lambda' = min\{\lambda, \beta\}$ και $\sigma' = min\{\sigma, \beta\}$
Θέτοντας $\lambda' = \sigma' = \beta$:
$r(\sigma, \lambda) = \frac{R}{1+\alpha} \cdot (\beta + \alpha\beta) = \frac{R}{1+\alpha}\beta(1+\alpha) = R\beta = \frac{R}{k}$, το οποίο είναι και αυτό που θέλουμε από το σχήμα μας.



Κατανομή stake με την αφελή προσέγγιση



Κατανομή stake με το [5]

References

1: Nakamoto, S. (2008). *Bitcoin: A Peer-to-Peer Electronic Cash System*. Retrieved from https://bitcoin.org/bitcoin.pdf

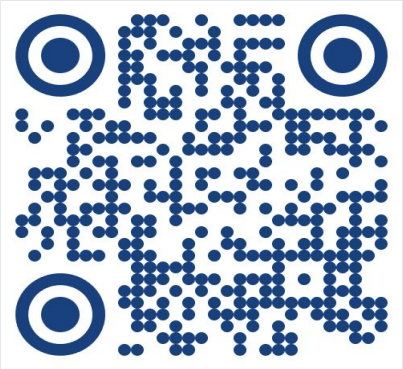
2: Wood, G. (2014). *Ethereum: A Secure Decentralised Generalised Transaction Ledger*. Ethereum Project Yellow Paper. Retrieved from https://ethereum.github.io/yellowpaper/paper.pdf

3: Easley, D., et al. (2020). *Flash Boys 2.0: Frontrunning, Transaction Fees, and Market Efficiency*. Review of Financial Studies. Retrieved from https://arxiv.org/pdf/1904.05234

4: Hoskinson, C., et al. (2017). *Cardano: A blockchain platform for the future*. Input Output Hong Kong. Retrieved from https://cardano.org

5: Brünjes, L., Kiayias, A., Koutsoupias, E., & Stouka, A. (2020). *Reward sharing schemes for stake pools*. https://arxiv.org/ftp/arxiv/papers/1807/1807.11218.pdf

THANKS



Find all my socials here – just scan!

CREDITS: This presentation template was created by **Slidesgo**, including icons by **Flaticon**, infographics & images by **Freepik**