



National Technical University of Athens

School of Electrical & Computer Engineering

Python Stock Market Analysis & Simulation

Konstantinos Poulinakis, MSc Student of DSML

AM : 03400153 , poulinakis.kon@gmail.com

Project completed as part of the course :

Programming Tools and Technologies for Data Science

February 27, 2022

Εισαγωγή

Σκοπός της παρούσας εργασίας είναι να δημιουργηθεί ένα σύστημα προσομοίωσης χρηματιστηριακών συναλλαγών και η επίτευξη υψηλού κέρδους μέσω της ανάπτυξης στρατηγικής που βασίζεται σε ανάλυση δεδομένων και η οποία υπόκειται σε ορισμένους περιορισμούς. Τα δεδομένα του χρηματιστηρίου από το 1960 και έπειτα θεωρούνται γνωστά και μπορούν να βρεθούν στον ακόλουθο [σύνδεσμο](#).

Μετά από ανάλυση των δεδομένων, αναπτύχθηκε μία στρατηγική η οποία έχει την δυνατότητα να επιφέρει πολύ υψηλά κέρδη τόσο για μία ακολουθία 1000 κινήσεων όσο και για ακολουθία με όριο το 1 εκατομμύριο κινήσεις.

Τα κέρδη που επιτυγχάνονται είναι της τάξεως των 7 δις για την ακολουθία 1000 κινήσεων και 62.5 δις για την μεγάλη ακολουθία .

Βασικά τμήματα Προσομοίωσης

Ο κώδικας έχει βασιστεί στην λογική του αντικειμενοστρεφούς προγραμματισμού. Το σκηνικό που προσπαθούμε να στήσουμε αποτελείται από 2 βασικές δομές οι οποίες περιγράφονται μέσω κλάσεων. Αυτές είναι ένας τραπεζικός λογαριασμός και ένα portfolio μετοχών.

Η κλάση του τραπεζικού λογαριασμού διαθέτει δύο μεθόδους, την ανάληψη και την κατάθεση χρημάτων. Οι μέθοδοι αυτοί χρησιμοποιούνται κάθε φορά που επιθυμούμε να εκτελέσουμε μια συναλλαγή στο χρηματιστήριο. Πριν από κάθε ανανέωση του διαθέσιμου ποσού εκτελούνται όλοι οι απαραίτητοι έλεγχοι όπως το επαρκές υπόλοιπο ή η κατάθεση θετικού ποσού.

Η άλλη απαραίτητη κλάση είναι αυτή του χαρτοφυλακίου. Αυτή περιέχει μεθόδους αφαίρεσης, εισαγωγής, εκτίμησης συνολικής αξίας και καταγραφής

του ιστορικού συναλλαγών. Πέρα από τις δύο αυτές κλάσεις έχει δημιουργηθεί η συνάρτηση `trade()` η οποία αναλαμβάνει να διεκπερεώσει και να ελέγξει για την ορθότητα τους όλες τις συναλλαγές που επιθυμούμε να εκτελέσουμε. Όλα αυτά συνιστούν το αρχείο `lib_stock.py`.

Η προσομοίωση του χρόνου και η εκτέλεση της στρατηγικής βρίσκεται στο αρχείο `time_travel.py`. Στο αρχείο αυτό πέρα από την στρατηγική που θα επεξηγηθεί αργότερα τρέχει ο βασικός βρόχος, μία επανάληψη για κάθε ημέρα μέχρι και την 2017-11-10.

Στρατηγική

Η στρατηγική που ακολουθήθηκε είναι βασισμένη έντονα στο Intraday Trading, δηλαδή την αγοροπωλησία μετοχών εντός της ίδιας ημέρας, για παραγωγή κέρδους. Η κύρια ιδέα είναι η εξής, προσπαθούμε να μαζέψουμε έναν τεράστιο αριθμό μετοχών μίας εταιρίας από την ημέρα ένταξης της στο χρηματιστήριο μέχρι και τις τελευταίες διαθέσιμες ημερομηνίες. Αφού έχουμε συγκεντρώσει έναν πολύ μεγάλο αριθμό μετοχών ξεκινάμε μαζική ρευστοποίηση των μετοχών μας περίπου 1 χρόνο πριν όταν η τιμή βρίσκεται σε πολύ υψηλά επίπεδα σε σχέση με όταν την πρωτοαγοράσαμε. Σκοπός είναι την τελευταία ημέρα όλες μας οι μετοχές να έχουν μετατραπεί σε ρευστό.

Φάσεις 1,2: Intra day Trading, Stock Accumulation

Κρίνεται σκόπιμο να μην πουλάμε μεγάλα ποσά κατά το πέραςμα του χρόνου καθώς το κέρδος τότε θα είναι πολύ χαμηλότερο. Πουλάμε ανά σχετικά σπάνια διαστήματα και μόνο εφόσον η τιμή της μετοχής βρίσκεται πάνω από ένα, μεταβαλλόμενο με τον χρόνο, κατώφλι (sell threshold). Για να μπορούμε να συλλέξουμε περισσότερες μετοχές προσπαθούμε να βγάζουμε αρκετά χρήματα εκμεταλευόμενοι την διακύμανση της τιμής μίας μετοχής

μέσα στην ίδια ημέρα. Έτσι όταν εντοπίσουμε ότι μία μετοχή πρόκειται να πέσει μέσα στην ίδια ημέρα τότε φροντίζουμε να πουλήσουμε σε μία υψηλή τιμή όσον το δυνατόν περισσότερες μετοχές μπορούμε (σεβόμενοι τα όρια που έχουν τεθεί από τον όγκο αγοροπωλησιών της μετοχής αυτής) με σκοπό να τις ξανά αγοράσουμε όλες σε χαμηλότερη τιμή την ίδια ημέρα. Αυτό μας δίνει το κέρδος της διαφοράς το οποίο μπορεί την επόμενη ημέρα να χρησιμοποιηθεί για να αγοραστούν ακόμα περισσότερες μετοχές. Όσο μεγαλύτερη είναι η διακύμανση των τιμών και όσο μεγαλύτερο όγκο μετοχών διαθέτουμε για να εκτελέσουμε αυτήν την αγοραπωλησία τόσο περισσότερο επωφελούμαστε από το intra day trading.

Πιο συγκεκριμένα οι εντολές που εκτελούμε είναι

- sell-high → buy-close
- sell-open → buy-low

Προφανώς είναι δυνατόν να εκτελέσουμε μόνο μία εκ των δύο εντολών, για αυτό για μεγιστοποίηση του κέρδους υπολογίζουμε εκ των προτέρων ποια από τις δύο εντολές πρόκειται να είναι η πιο επικερδής βασιζόμενοι στην διαφορά των τιμών high, low, open, close.

Με άλλα λόγια κάθε φορά που η υπό εξέταση μετοχή πρόκειται να δεχθεί ένα πλήγμα στην τιμή της εμείς όχι απλά δεν χάνουμε λεφτά αλλά επωφελούμαστε έντονα από αυτό, σαν να μπορούσαμε να προβλέψουμε το μέλλον !

Προφανώς βέβαια για να καταφέρουμε να χτίσουμε επαρκές κεφάλαιο είναι αναγκαίο να εκτελούμε και κάποιες κανονικές πωλήσεις. Για αυτό τον λόγο κάθε ημέρα ο πράκτορας (agent) καλείται να λάβει μία απόφαση. Εφόσον η μεταβολή τιμών δεν είναι αρκετά μεγάλη ώστε να εκτελέσουμε intra day trading τότε ο πράκτορας θα λάβει μία απόφαση από τις εξής τρεις.

Action	Probability
Buy	0.65
Hold	0.125
Sell	0.125

Table 1: Πιθανότητες επιλογής δράσης.

Οι παραπάνω τιμές έχουν στην λογική τους ότι θέλουμε συχνότερα να αγοράζουμε και σπανιότερα να πουλάμε, αλλά οι ακριβείς τιμές έχουν προκύψει πειραματικά.

Πέρα από την στοχαστική φύση της επιλογής, υπάρχουν και κάποιοι κανόνες που θα πρέπει να ισχύουν για να εκτελεστεί κάποια δράση. Συγκεκριμένα έχουμε ορίσει δύο όρια, τα buy-threshold και sell-threshold. Όσες μετοχές βρίσκονται κάτω από το όριο buy-threshold γίνονται υποψήφιες για αγορά, ενώ όσες μετοχές βρίσκονται πάνω από την τιμή sell-threshold γίνονται υποψήφιες για πώληση. Τα κατώφλια αυτά είναι σημαντικό να έχουν διαφορά μεταξύ τους και το sell-threshold να είναι υψηλότερο από το buy, διαφορετικά είναι πιθανό να εκτελούμε εντολές που μας οδηγούν σε ζημία.

$$Buy = \begin{cases} Yes & \text{if } Price \leq buy_threshold \\ No & \text{if } Price > buy_threshold \end{cases} \quad (1)$$

$$Sell = \begin{cases} Yes & \text{if } Price \geq sell_threshold \\ No & \text{if } Price < sell_threshold \end{cases} \quad (2)$$

Ανάλογα με την εφαρμογή οι αρχικές τιμές κατωφλίων ποικίλουν, είναι όμως τέτοιες ώστε να μπορούμε να ξεκινήσουμε να αγοράζουμε χωρίς πρόβλημα μία μετοχή αλλά να την πουλάμε μόνο εφόσον η τιμή έχει αυξηθεί σε σχέση με την αρχική αγορά. Για αυτό στο τέλος κάθε ημέρας τα κατώφλια αυξάνονται κατά έναν παράγοντα ώστε να λαμβάνεται υπόψιν η γενική αύξηση της αξίας των μετοχών στο χρηματιστήριο με την πάροδο του χρόνου. Ο παράγοντας μεταβολής είναι μία υπερπαράμετρος της στρατηγικής. Να αναφερθεί ότι

δοκιμάστηκαν γραμμικοί, λογαριθμικοί και εκθετικοί παράγοντες. Λογαριθμικοί παράγοντες χαμηλών τιμών πχ. $\log(5)$, $\log(4)$ φάνηκε να αποφέρουν αρκετά καλά αποτελέσματα. Αυτά τα κατώφλια είναι ιδιαίτερα σημαντικά κατά τα πρώτα έτη και σιγά σιγά χάνουν την αξία τους όσο το intra day trading γίνεται πιο σημαντικό.

$$BUY_THRESHOLD_{t+1} = b \times BUY_THRESHOLD_t$$

$$SELL_THRESHOLD_{t+1} = s \times SELL_THRESHOLD_t$$

Φάση 3: Margin Call !

Στην τρίτη και τελευταία φάση της στρατηγικής γίνεται προσπάθεια να ρευστοποιηθεί ολόκληρο το χαρτοφυλάκιο και η αξία των μετοχών να μετατραπεί σε ρευστό. Η τρίτη φάση πυροδοτείται με την άφιξη κάποιας συγκεκριμένης ημερομηνίας (liquidation_day). Από την liquidation_day και έπειτα εκτελούνται μόνο πωλήσεις μετοχών. Οι αγορές μετοχών παγώνουν ενώ το intra day trading κατά τις ημέρες εκείνες παρέχει πολύ μικρότερο κέρδος διότι ο διαθέσιμος όγκος αγοραπωλησιών καταναλώνεται κατά την πώληση. Ως εκ τούτου, τουλάχιστον για τον περιορισμό των 1000 κινήσεων, είναι χρήσιμο να παγωθεί και το intra day trade κατά την φάση ρευστοποίησης.

Η επιλογή της liquidation_day, αποτελεί υπερπαράμετρο της λύσης και η επιλογή γίνεται πειραματικά. Να αναφερθεί ότι εάν η ημερομηνία αυτή έρθει πολύ αργά τότε είναι πολύ πιθανό ότι δεν θα προλάβουμε να ρευστοποιήσουμε όλες μας τις μετοχές και το χρήμα θα μείνει "παγιδευμένο". Αντίθετα αν η ημερομηνία έρθει πολύ νωρίς τότε πρακτικά χάνουμε κέρδη από την φάση 2, αφού καταναλώνουμε λιγότερο χρόνο και κινήσεις σε αγορά μετοχών και intra day trading. Μέσω πειραματισμού βρέθηκαν σχεδόν βέλτιστες ημερομηνίες και για τις δύο ακολουθίες κινήσεων.

Ανάλυση και επιλογή Μετοχών

Εφόσον έχουμε αναπτύξει την στρατηγική μας, έχει έρθει η ώρα να κάνουμε μία επιλογή των κατάλληλων μετοχών με τις οποίες θα ενορχηστρώσουμε το χαρτοφυλάκιο μας. Στα δεδομένα που διαθέτουμε έχουμε πληροφορίες για τις τιμές open, close, low, high και για τον όγκο ανταλλαγής μια μετοχής, volume.

Growth - Ratio

Για να αποφέρει καρπούς η στρατηγική μας είναι πολύ σημαντικό μία μετοχή να έχει πολύ καλή ανάπτυξη κατά την πάροδο των ετών. Έτσι ορίζοντας τον δείκτη

$$growth_ratio = \frac{initial_price}{final_price},$$

όπου initial_price η τιμή της μετοχής κατά την ημέρα του IPO (1η ημέρα στο χρηματιστήριο) και final_price η τιμή της μετοχής κατά την τελευταία διαθέσιμη ημερομηνία που υπάρχει στα δεδομένα μας (2017-11-10). Θέτοντας ένα κάτω όριο ίσο με 400 και αφού έχουμε καθαρίσει το dataset από λανθασμένες μηδενικές τιμές έναρξης λαμβάνουμε την παρακάτω λίστα, εικ. 2. Στην πρώτη στήλη αναγράφεται ο κωδικός της εταιρίας, στην 2η το growth ratio που ορίσαμε παραπάνω και στην 3η η ημερομηνία ένταξης της μετοχής στο χρηματιστήριο.

```
[12]: # Sort the List of growth companys by growth ratio
growth.sort(key= lambda x : x[1], reverse=True)
growth

[12]: [
    ['hd', 6318.76, '1981-09-22'],
    ['mo', 2935.62, '1970-01-02'],
    ['intc', 2891.96, '1972-01-07'],
    ['wmt', 2651.78, '1972-03-20'],
    ['amgn', 1975.7, '1984-09-07'],
    ['msft', 1246.88, '1986-03-13'],
    ['brk-a', 949.24, '1980-03-17'],
    ['uve', 907.95, '2005-02-25'],
    ['adbe', 861.9, '1986-08-14'],
    ['unh', 809.39, '1990-03-26'],
    ['mcd', 587.2, '1970-01-02'],
    ['amzn', 571.62, '1997-05-16'],
    ['csc', 516.14, '1990-03-26'],
    ['gd', 482.47, '1977-01-03'],
    ['ben', 474.33, '1984-09-07'],
    ['twx', 444.91, '1992-03-19'],
    ['lmt', 420.01, '1977-01-03'],
    ['aapl', 413.11, '1984-09-07'],
    ['bby', 407.68, '1985-04-19'],
```

Figure 1: Εταιρίες με την μεγαλύτερη ανάπτυξη τιμής από την ένταξη τους στο χρηματιστήριο μέχρι και την 2017-11-10.

Φθηνή αρχική τιμή

Λόγω των περιορισμών που έχουμε, κατά την πρώτη μας ημέρα στο χρηματιστήριο διαθέτουμε μονάχα 1 δολλάριο στον λογαριασμό μας, συνεπώς είναι απαραίτητο να φιλτράρουμε τις εταιρίες αυτές ώστε να απομείνουν μόνο όσες ξεκινάνε με φθηνή τιμή πχ. μικρότερη του μισού δολλαρίου 0.5\$. Ευτυχώς, λόγω της ημερομηνίας ένταξης τους οι περισσότερες από αυτές ξεκινούν με τιμές κάτω από 0.1\$.

Μεγάλος όγκος συναλλαγών

Μία τρίτη άκρως σημαντική παράμετρος για την επιτυχία της στρατηγικής μας είναι η μετοχή αυτή να έχει **πολύ μεγάλο όγκο συναλλαγών** καθώς έτσι μεγιστοποιείται το κέρδος από το intra day, μπορούμε να αγοράζουμε περισσότερες μετοχές κάθε ημέρα αλλά και να τις ρευστοποιήσουμε γρήγορα

στην 3η φάση. Συνεπώς εκτελούμε ακόμα ένα φιλτράρισμα για να βρούμε τις μετοχές με τον μεγαλύτερο μέσο όγκο συναλλαγών (mean volume) καθόλη την διάρκεια των ετών. Η ανάλυση αυτή μας δίνει την παρακάτω λίστα

```
# Sort by volume
volume.sort(key=lambda x : x[1], reverse=True)
volume

[['aapl', 106641628.3860593],
 ['msft', 79458003.26218213],
 ['csc', 71469592.13670304],
 ['intc', 56466008.66528211],
 ['bac', 54419994.619245805],
 ['siri', 49223705.18525461],
 ['orcl', 44536613.045818865],
 ['ebay', 41298413.496260904],
 ['fb', 37707164.986241855],
 ['s', 36728107.33],
 ['dd', 34861021.0],
 ['mo', 32593646.064518802],
 ['t', 31099863.08975122],
 ['vale', 28418644.384879727],
 ['brk-b', 27890478.004062787],
 ['sogo', 27541219.5],
 ['snap', 26809068.38202247],
 ['pfe', 26292679.951669984],
 ['nok', 26279031.153389566]]
```

Figure 2: Εταιρίες με τον μεγαλύτερο μέσο όγκο συναλλαγών (αριθμός αγοραπωλησιών μετοχής ανά ημέρα)

Μετά από την ανάλυση αυτή μπορούμε να πούμε ότι πολλά υποσχόμενες μετοχές, χωρίς συγκεκριμένη σειρά είναι οι :

- AAPL (apple)
- MSFT (microsoft)
- INTC (intel)

Παρουσίαση Αποτελεσμάτων

Μετά από αρκετό πειραματισμό κατέληξα ότι το πιο κερδοφόρο ήταν να επιλέξω μόνο μία από τις 3 παραπάνω μετοχές. Αρχικά η επιλογή ήταν η

εταιρία Intel καθώς είναι η παλαιότερη, κάτι που θεωρητικά μου επιτρέπει περισσότερες κινήσεις και καθώς είχε το μεγαλύτερο growth factor. Τα αποτελέσματα με την Intel ήταν αρκετά ικανοποιητικά με κέρδος που ξεπερνούσε τα 5 δις δολάρια. Στην συνέχεια δοκιμάστηκε η μετοχή Apple κυρίως λόγω του τεράστιου όγκου συναλλαγών της. Όμως δεν απέδωσε όπως θα περιμέναμε, κυρίως λόγω της αυξημένης αρχικής της τιμής, περίπου 0.45\$. Πιθανότατα να μην βρέθηκε ο κατάλληλος συνδυασμός παραμέτρων ώστε να εκτοξευθεί ο αριθμός μετοχών. Έτσι τέλος δοκιμάστηκε η εταιρία Microsoft, η οποία φέροντας τα θετικά και των δύο προηγούμενων, χαμηλή τιμή, μεγάλος όγκος και υψηλό growth, κατάφερε να αποδώσει περίφημα.

Τα αποτελέσματα με την εταιρία Microsoft ήταν κέρδη της τάξεως των 8 δις\$ στην ακολουθία 1000 κινήσεων και κέρδη της τάξεως των 62 δις\$ στην μεγάλη ακολουθία με συνολικό αριθμό κινήσεων 14041. Λόγω της στοχαστικής φύσης της στρατηγικής τα αποτελέσματα διαφέρουν αλλά κινούνται πολύ κοντά σε αυτά που αναφέρθηκαν εφόσον οι παράμετροι μείνουν ως έχουν. Παρακάτω παρουσιάζονται τα αποτελέσματα και τα διαγράμματα της στρατηγικής που παρουσιάστηκε, χρησιμοποιώντας την εταιρία Microsoft.

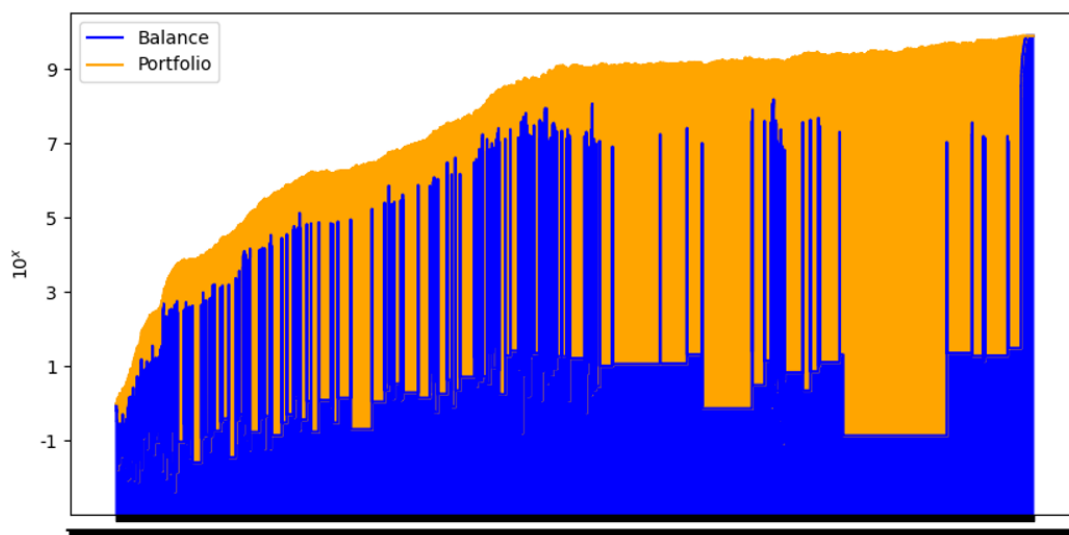
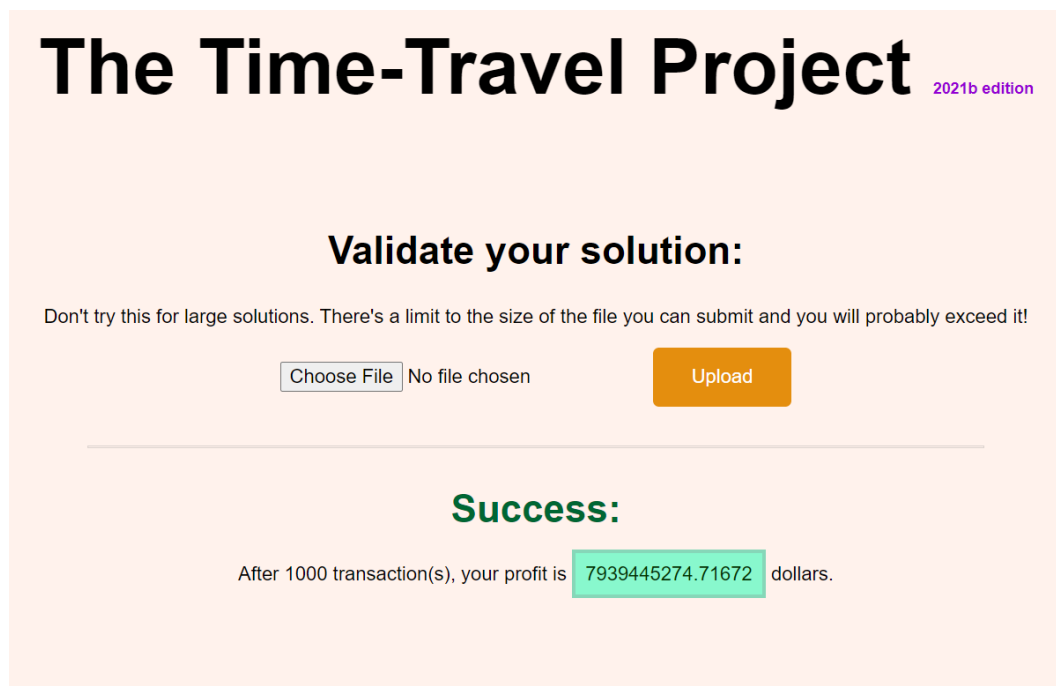


Figure 3: Γράφημα αξίας χαρτοφυλακίου και τραπεζικού λογαριασμού. Ακολουθία 1000 κινήσεων.



The Time-Travel Project

2021b edition

Validate your solution:

Don't try this for large solutions. There's a limit to the size of the file you can submit and you will probably exceed it!

Choose File No file chosen Upload

Success:

After 1000 transaction(s), your profit is 7939445274.71672 dollars.

Figure 4: Το κέρδος της ακολουθίας 1000 κινήσεων ξεπερνά τα 6 δισεκατομμύρια.

Στον οριζόντιο άξονα βρίσκεται ο χρόνος, ενώ στον κατακόρυφο έχουμε αξία σε λογαριθμική κλίμακα 10^x .

Παρατηρήστε πώς κατά τις τελευταίες ημέρες του γραφήματος όλη η αξία του χαρτοφυλακίου (μετοχές + λογαριασμός), μετατρέπεται σε ρευστό με πολύ ταχύ ρυθμό λόγω της ρευστοποίησης.

Παρατηρήστε επίσης τις ραγδαίες διακυμάνσεις στο υπόλοιπο του λογαριασμού (μπλε γραμμές), αυτές οι διακυμάνσεις συμβαίνουν εξαιτίας της πολύ επιθετικής αγοραπωλησίας που ακολουθείται ως μέρος της στρατηγικής.

Τα ίδια, μπορούν να παρατηρηθούν και στο παρακάτω γράφημα για την μεγάλη ακολουθία.

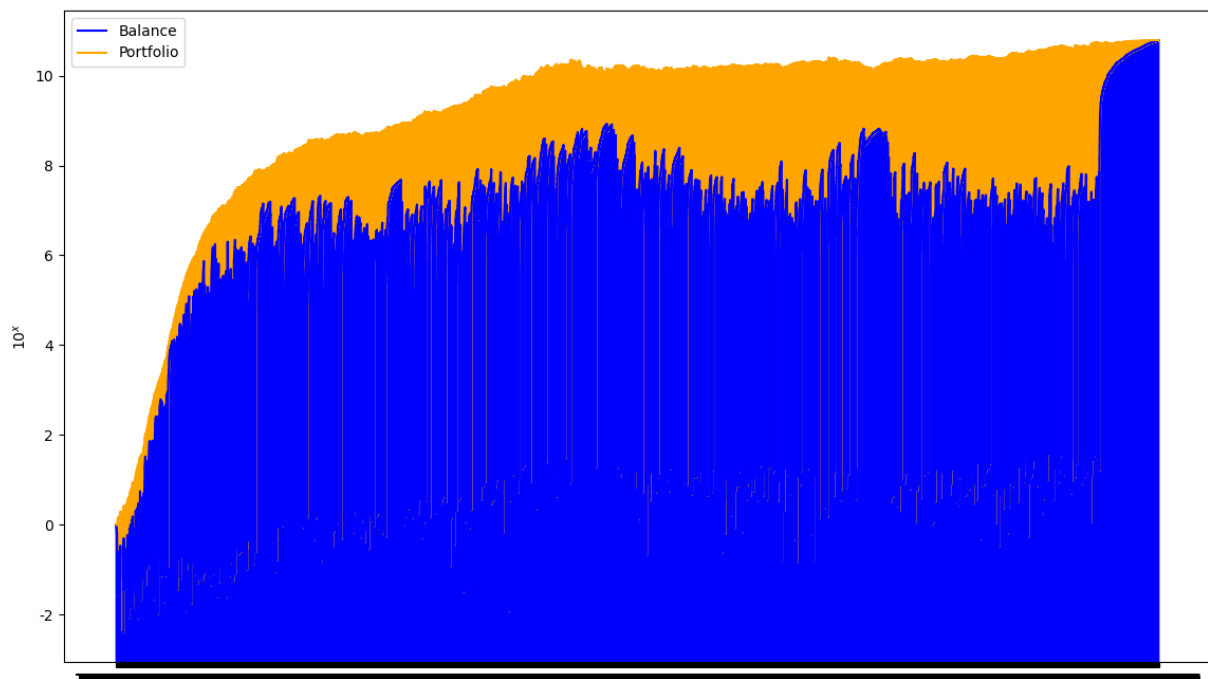


Figure 5: Γράφημα αξίας χαρτοφυλακίου και τραπεζικού λογαριασμού. Ακολουθία 14041 κινήσεων.

The Time-Travel Project

2021b edition

Validate your solution:

Don't try this for large solutions. There's a limit to the size of the file you can submit and you will probably exceed it!

No file chosen

Success:

After 14041 transaction(s), your profit is 62451780257.794975 dollars.

Figure 6: Το κέρδος της ακολουθίας 14041 κινήσεων ξεπερνά τα 62 δισεκατομμύρια.