

Project HY240b - 1η Φάση

Διδάσκουσα: Παναγιώτα Φατούρου

Βοηθός: Νικόλαος Μπατσαράς

Τμήμα Επιστήμης Υπολογιστών
Πανεπιστήμιο Κρήτης

Ηράκλειο, 09 Μαρτίου 2018



Εισαγωγή

Διαδικαστικά
Θεματολογία

Περιγραφή Δομών

Γενικά
Δομές Χρηστών
Δομές Tweets

Γεγονότα

Γενικά
Είδη Γεγονότων

Εισαγωγή

Διαδικαστικά

Θεματολογία

Περιγραφή Δομών

Γενικά

Δομές Χρηστών

Δομές Tweets

Γεγονότα

Γενικά

Είδη Γεγονότων

Πληροφορίες

- ▶ Σελίδα Μαθήματος: <https://csd.uoc.gr/~hy240b>
- ▶ Τρόπος Παράδοσης: Με το πρόγραμμα turnin
 - ▶ Περισσότερες Πληροφορίες:
<https://csd.uoc.gr/~hy240b/current/submit.php>
- ▶ Παράδοση: Δευτέρα, 01 Απριλίου 2018, ώρα 23:59
- ▶ Διαθέσιμες Υλοποιήσεις: C ή Java
- ▶ Ερωτήσεις-Απορίες: hy240b-list@csd.uoc.gr
- ▶ **Προσοχή:** Compile/Run στα μηχανήματα της σχολής.

Εισαγωγή

Διαδικαστικά

Θεματολογία

Περιγραφή Δομών

Γενικά

Δομές Χρηστών

Δομές Tweets

Γεγονότα

Γενικά

Είδη Γεγονότων

Θέμα

Προσομοίωση ενός μέσου κοινωνικής δικτύωσης παρόμοιο με το Twitter.

- ▶ Χρήστες που στέλνουν και διαβάζουν σύντομα μηνύματα, τα οποία ονομάζονται τουίτς (tweets)
- ▶ Ο κάθε χρήστης μπορεί να “ακολουθεί” (follow) άλλους χρήστες με σκοπό να εμφανίζονται τα tweets τους στην περιοχή του (Wall)

Θέμα

Έχουμε λοιπόν 2 βασικές οντότητες:

1. Χρήστες
2. Tweets

Εισαγωγή

Διαδικαστικά
Θεματολογία

Περιγραφή Δομών

Γενικά
Δομές Χρηστών
Δομές Tweets

Γεγονότα

Γενικά
Είδη Γεγονότων

Γνωστικό Αντικείμενο

Οι γνώσεις που θα χρειαστούν για την υλοποίηση της 1ης φάσης της εργασίας, μεταξύ άλλων, είναι οι εξής:

- ▶ Λίστες
 - ▶ Αναζήτηση, Εισαγωγή, Διαγραφή, Συγχώνευση
 - ▶ Ταξινομημένες λίστες
 - ▶ Κόμβοι φρουροί

Ιδέα

Η εργασία αποτελείται από 2 ομάδες δομών:

1. Δομές που αφορούν τους χρήστες
2. Δομές που αφορούν τα tweets

Εισαγωγή

Διαδικαστικά
Θεματολογία

Περιγραφή Δομών

Γενικά
Δομές Χρηστών
Δομές Tweets

Γεγονότα

Γενικά
Είδη Γεγονότων

Γενικά

Οι δομές χρηστών οργανώνονται ως εξής:

- ▶ Οι χρήστες αποθηκεύονται σε μία **απλά συνδεδεμένη, ταξινομημένη σε αύξουσα διάταξη** λίστα
- ▶ Ο κάθε χρήστης είναι ένα αντικείμενο τύπου struct user
- ▶ Κάθε χρήστης περιέχει ένα δείκτη στην αρχή μίας **απλά συνδεδεμένης** λίστας, τύπου struct follower, της οποίας κάθε κόμβος συμβολίζει έναν ακόλουθο του χρήστη αυτού
- ▶ Κάθε χρήστης περιέχει επίσης, κι ένα δείκτη στην αρχή μίας **απλά συνδεδεμένης** λίστας με **κόμβο φρουρό**, τύπου struct tweet_w, της οποίας κάθε κόμβος συμβολίζει ένα tweet στην περιοχή (Wall) του χρήστη αυτού

Δομή Χρηστών

Χρήστης - struct user:

- uid** Αναγνωριστικό χρήστη (int)
- followers** Δείκτης στη αρχή μίας λίστας, της οποίας ο κάθε κόμβος αποτελεί έναν follower του χρήστη αυτού (struct follower *)
- tweets** Δείκτης στην αρχή μίας λίστας, της οποίας ο κάθε κόμβος αποτελεί ένα tweet στην περιοχή του χρήστη αυτού (struct tweet_w *)
- next** Δείκτης στον επόμενο κόμβο στην λίστα χρηστών (struct user *)

Δομή Οπαδών

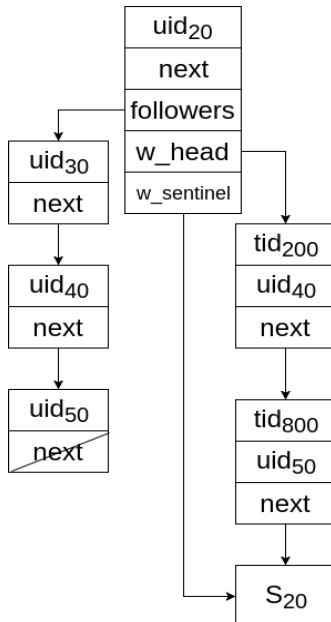
Follower - struct follower:

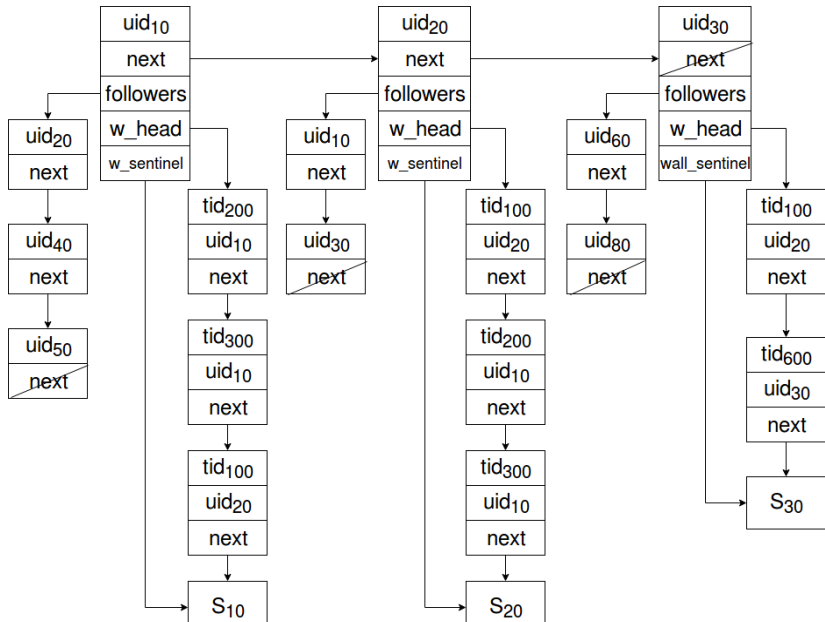
- uid** Αναγνωριστικό χρήστη/οπαδού (int)
- next** Δείκτης στον επόμενο κόμβο στην λίστα οπαδών (struct follower *)

Δομή Tweets

Tweet - struct tweet_w:

- tid** Αναγνωριστικό tweet (int)
- uid** Αναγνωριστικό χρήστη που δημοσίευσε το tweet (int)
- next** Δείκτης στον επόμενο κόμβο στην λίστα των tweets (struct tweet_w *)





Εισαγωγή

Διαδικαστικά
Θεματολογία

Περιγραφή Δομών

Γενικά
Δομές Χρηστών
Δομές Tweets

Γεγονότα

Γενικά
Είδη Γεγονότων

Γενικά

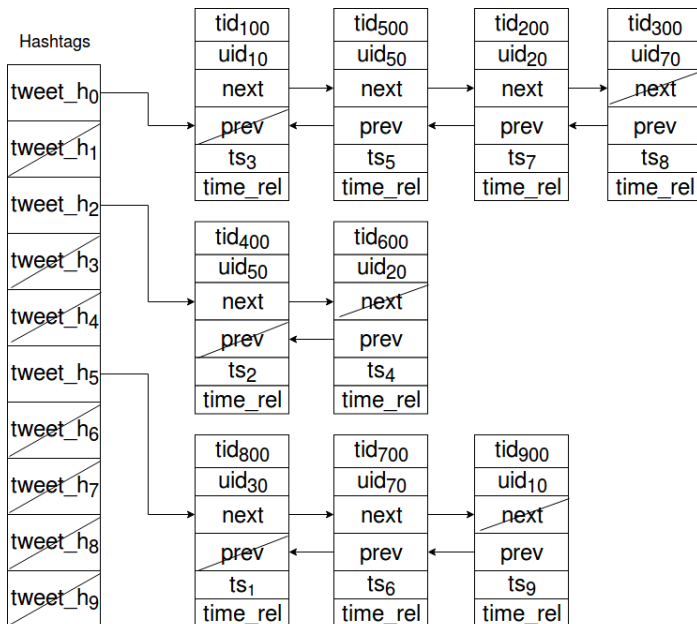
Οι δομές των tweets οργανώνονται ως εξής:

- ▶ Τα tweets αποθηκεύονται σε ένα **πίνακα**
- ▶ Κάθε θέση i του πίνακα είναι ένας δείκτης σε μία **διπλά συνδεδεμένη, ταξινομημένη με βάση το timestamp** λίστα που ονομάζεται λίστα tweets της κατηγορίας i και περιέχει όλα τα tweets με hashtag i
- ▶ Το κάθε tweet είναι ένα αντικείμενο τύπου struct tweet_h

Δομή Tweets

Tweet - struct tweet_h:

- tid** Αναγνωριστικό tweet (int)
- timestamp** Ακέραιος που αντιστοιχεί στην ημερομηνία δημοσίευσης του tweet (int)
- time_rel** Ακέραιος αριθμός που αντιστοιχεί στην χρονική ισχύ ενός tweet (int)
- next** Δείκτης στον επόμενο κόμβο στην λίστα των tweets (struct tweet_h *)
- prev** Δείκτης στον προηγούμενο κόμβο στην λίστα των tweets(struct tweet_h *)



Εισαγωγή

Διαδικαστικά
Θεματολογία

Περιγραφή Δομών

Γενικά
Δομές Χρηστών
Δομές Tweets

Γεγονότα

Γενικά
Είδη Γεγονότων

Ιδέα

Ένα γεγονός αποτελεί μία ενέργεια την οποία εκτελεί το σύστημά μας, π.χ.

- ▶ Εισαγωγή ενός χρήστη
- ▶ Η δημοσίευση ενός tweet
- ▶ κ.τ.λ.

Πηγή

- ▶ Τα γεγονότα αυτά είναι καταχωρημένα σε κάποια αρχεία
- ▶ Τα αρχεία υπάρχουν στη σελίδα του μαθήματος
- ▶ Το πρόγραμμα κάνει parse αυτά τα αρχεία και εκτελεί τις συνάρτηση που αντιστοιχεί στο εκάστοτε γεγονός
- ▶ Δεν χρειάζεται να γραφτεί επιπλέον κώδικας για το χειρισμό των γεγονότων
- ▶ Το μόνο που λείπει είναι η υλοποίηση των συναρτήσεων

Μορφή

Ένα γεγονός έχει την εξής μορφή:

```
event_id [arithmetic_arguments [...]]
```

Εισαγωγή

Διαδικαστικά
Θεματολογία

Περιγραφή Δομών

Γενικά
Δομές Χρηστών
Δομές Tweets

Γεγονότα

Γενικά
Είδη Γεγονότων

Register

Εισαγωγή ενός νέου χρήστη στο σύστημα.
Έχει την εξής μορφή:

R uid

Ορίσματα:

uid Αναγνωριστικό του νέου χρήστη

Register

Κατά το γεγονός αυτό:

1. Θα γίνει εισαγωγή ενός νέου χρήστη με αναγνωριστικό uid στη λίστα χρηστών
2. **Προσοχή**: Μετά από κάθε εισαγωγή, η λίστα χρηστών θα πρέπει να παραμένει ταξινομημένη

Subscribe

Εγγραφή του χρήστη με αναγνωριστικό uid_2 στη λίστα των followers του χρήστη με αναγνωριστικό uid_1 .

Έχει την εξής μορφή:

$$S\ uid_1\ uid_2$$

Ορίσματα:

uid_1 Αναγνωριστικό του χρήστη που έλαβε έναν νέο follower

uid_2 Αναγνωριστικό του χρήστη που κάνει follow

Subscribe

Κατά το γεγονός αυτό:

1. θα αναζητήσετε στη λίστα χρηστών το χρήστη με αναγνωριστικό uid_1
2. θα εισάγετε στη λίστα των followers ένα νέο κόμβο με αναγνωριστικό uid_2
3. **Προσοχή:** Μετά από κάθε εισαγωγή, η λίστα των followers θα πρέπει να παραμένει ταξινομημένη

Tweet

Δημοσίευση ενός tweet με αναγνωριστικό *tid* από το χρήστη με αναγνωριστικό *uid*.

Έχει την εξής μορφή:

T uid tid hashtag timestamp time_relevance

Ορίσματα:

uid Αναγνωριστικό του χρήστη που δημοσίευσε το tweet

tid Αναγνωριστικό του tweet

hashtag Κατηγορία στην οποία ανοίκει το tweet

timestamp Χρονική στιγμή την οποία δημοσιεύτηκε το tweet

time_rel Χρονική ισχύ του tweet

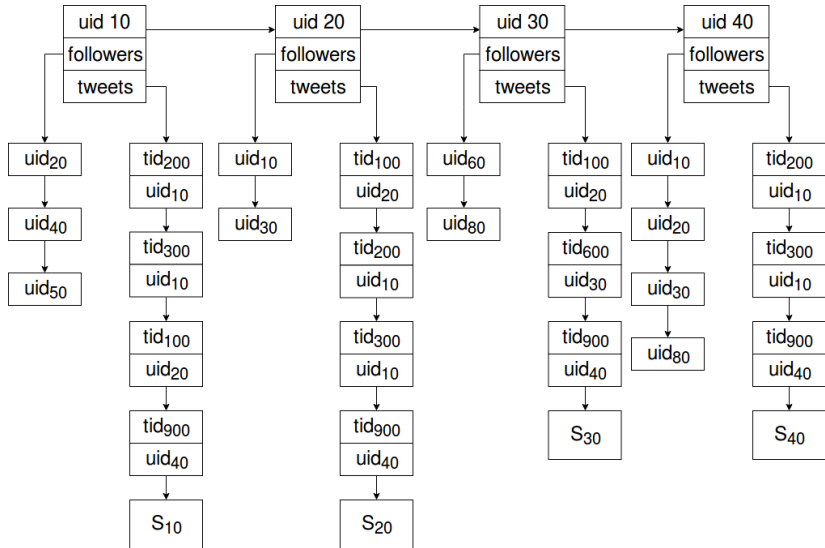
Tweet

Κατά το γεγονός αυτό:

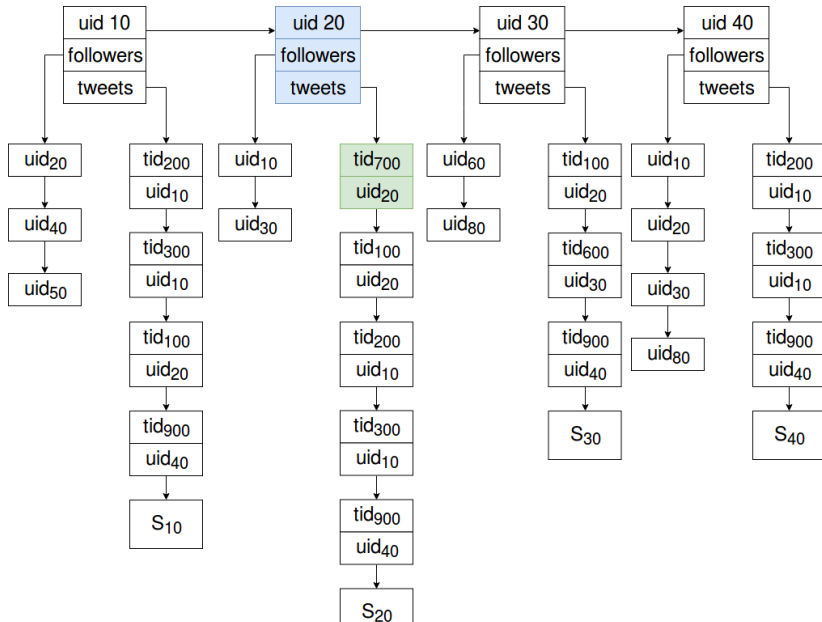
1. Θα αναζητήσετε στη λίστα χρηστών το χρήστη με αναγνωριστικό uid
2. Θα εισάγετε στη λίστα των tweets του, ένα νέο κόμβο με αναγνωριστικό tid
3. Θα διατρέξετε τη λίστα των followers του χρήστη αυτού και για κάθε κόμβο στη λίστα:
 - 3.1 Θα αναζητείτε το χρήστη με το τρέχον uid στη λίστα χρηστών
 - 3.2 Αφού τον εντοπίσετε, θα εισάγετε στη λίστα των tweets του, ένα νέο κόμβο με αναγνωριστικό tid
 - 3.3 **Προσοχή:** Η διαδικασία αυτή, θα πρέπει να εκτελείται σε χρόνο $O(n)$, όπου n είναι το πλήθος των κόμβων στη λίστα χρηστών
4. Θα εισάγετε και ένα νέο κόμβο στον πίνακα κατηγοριών

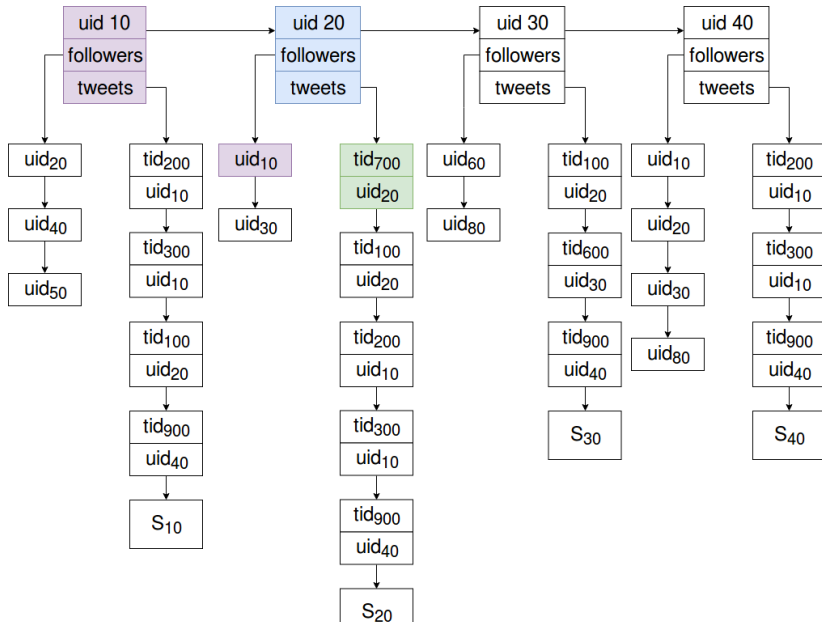
Tweet: Παράδειγμα

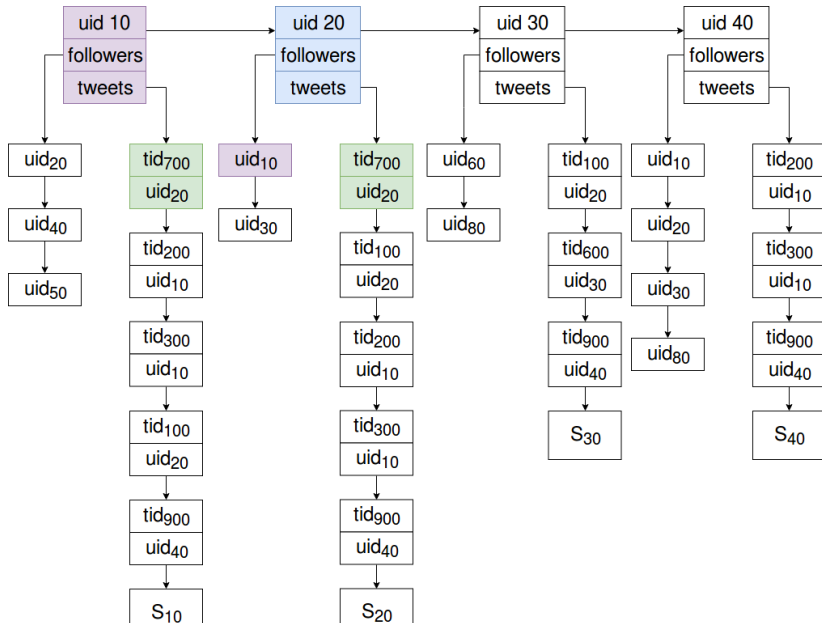
- ▶ Έστω γεγονός T 20 700 2 20180416 3



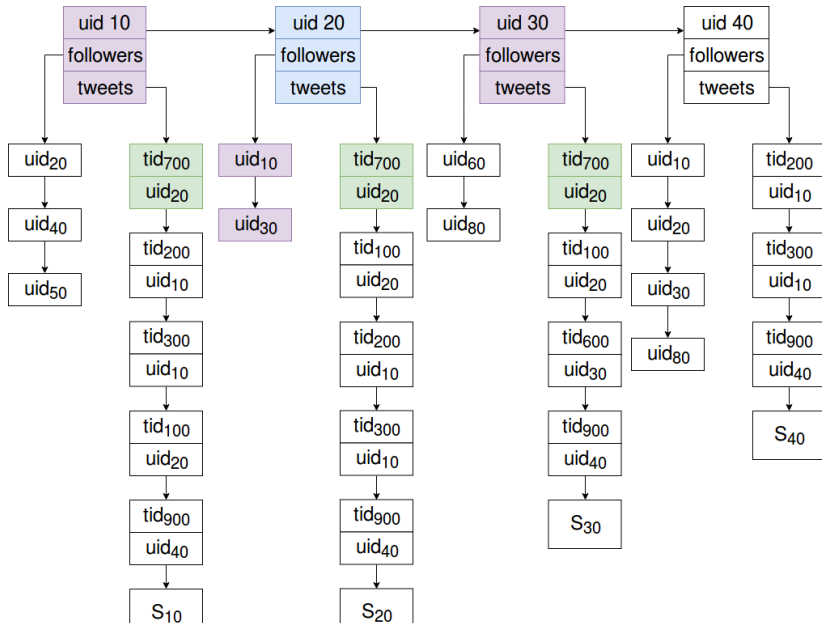












Unsubscribe

Διαγραφή του χρήστη με αναγνωριστικό uid_2 από τη λίστα των followers του χρήστη με αναγνωριστικό uid_1 .

Έχει την εξής μορφή:

$$U\ uid_1\ uid_2$$

Ορίσματα:

uid_1 Αναγνωριστικό του χρήστη που έχασε ένα follower

uid_2 Αναγνωριστικό του χρήστη που κάνει unsubscribe

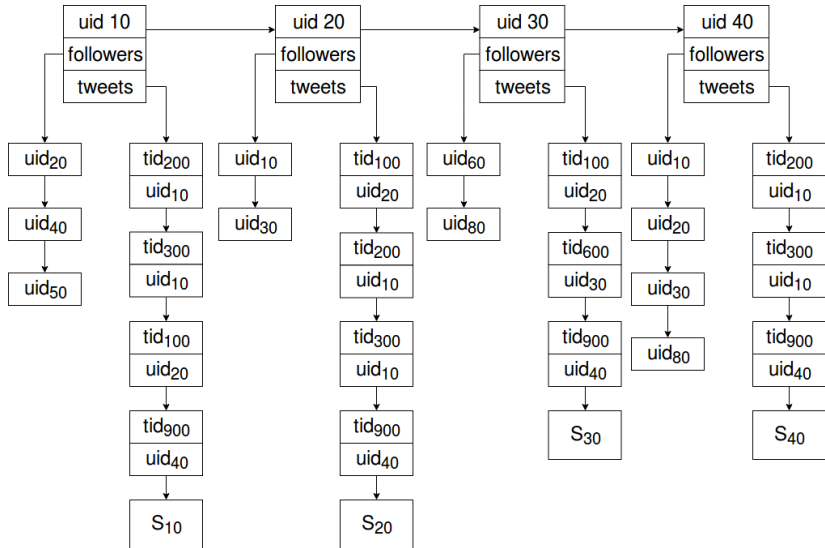
Unsubscribe

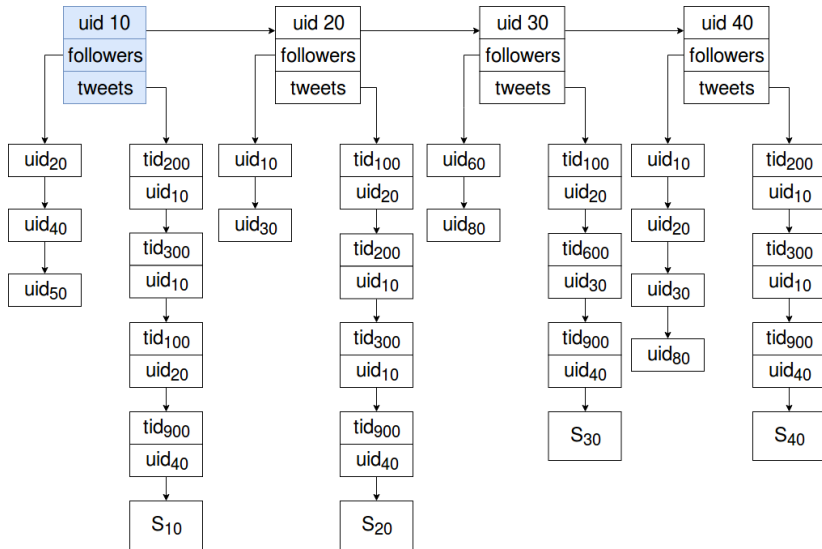
Κατά το γεγονός αυτό:

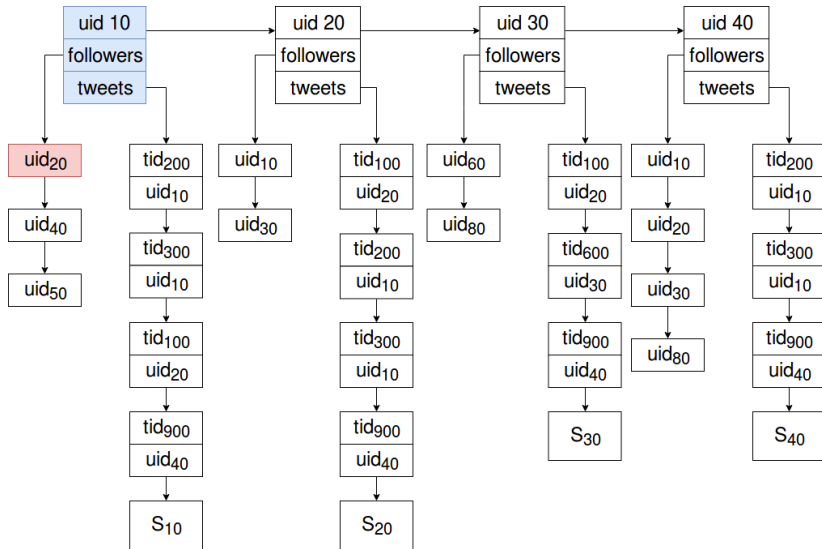
1. Θα αναζητήσετε στη λίστα χρηστών το χρήστη με αναγνωριστικό uid_1
2. Θα αφαιρέσετε από λίστα των followers τον κόμβο με αναγνωριστικό uid_2
3. Θα αναζητήσετε στη λίστα χρηστών το χρήστη με αναγνωριστικό uid_2
4. Θα διατρέξετε τη λίστα των tweets του χρήστη με αναγνωριστικό uid_2
5. Θα αφαιρέσετε από αυτή τους κόμβους που έχουν το πεδίο uid ίσο με uid_1

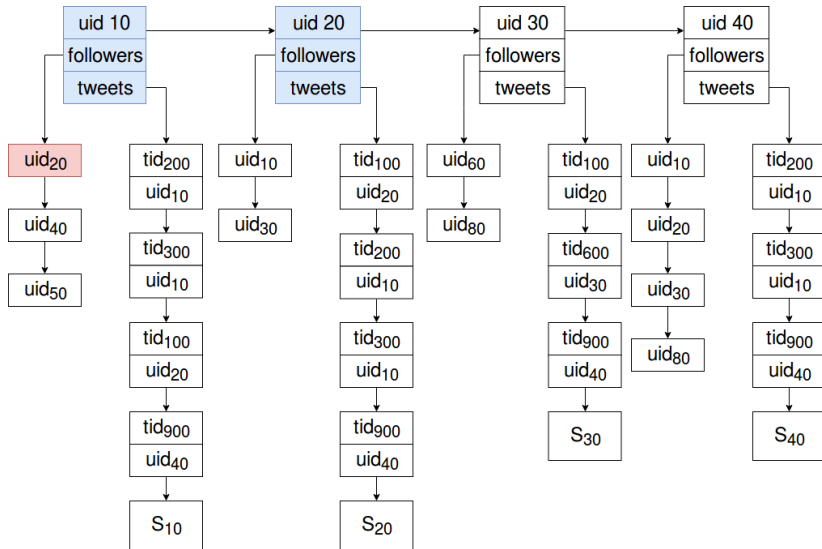
Unsubscribe: Παράδειγμα

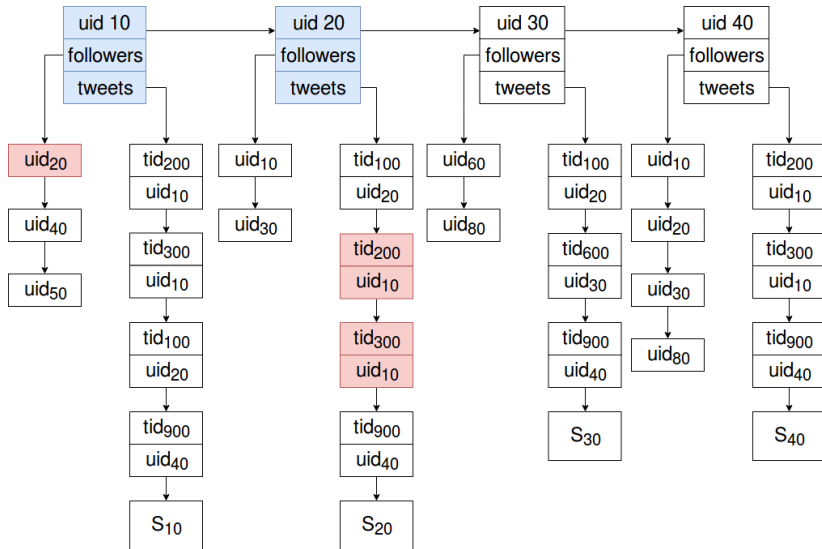
- ▶ Έστω γεγονός $U_{10\ 20}$











Delete

Διαγραφή του χρήστη με αναγνωριστικό *uid* από το σύστημα.
Έχει την εξής μορφή:

D uid

Ορίσματα:

uid Αναγνωριστικό του χρήστη προς διαγραφή

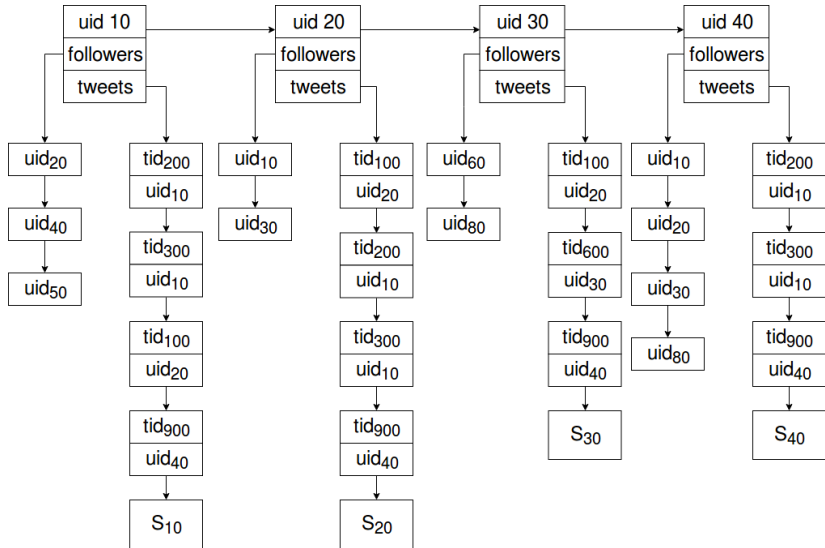
Delete

Κατά το γεγονός αυτό:

1. Θα αναζητήσετε στη λίστα χρηστών το χρήστη με αναγνωριστικό *uid*
2. Θα διαγράψετε όλα τα στοιχεία της λίστας tweets αυτού του χρήστη
3. Θα διατρέξετε τη λίστα των οπαδών και για κάθε κόμβο:
 - 3.1 Θα καλείτε κατάλληλα τη συνάρτηση που υλοποιεί το γεγονός unsubscribe
4. Θα διαγράφετε από λίστα χρηστών το χρήστη με αναγνωριστικό *uid*

Delete: Παράδειγμα

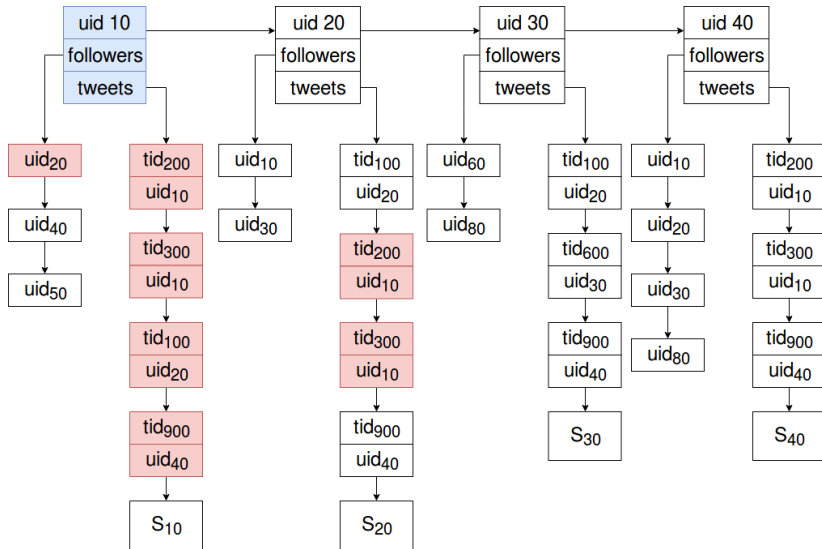
- ▶ Έστω γεγονός D 10

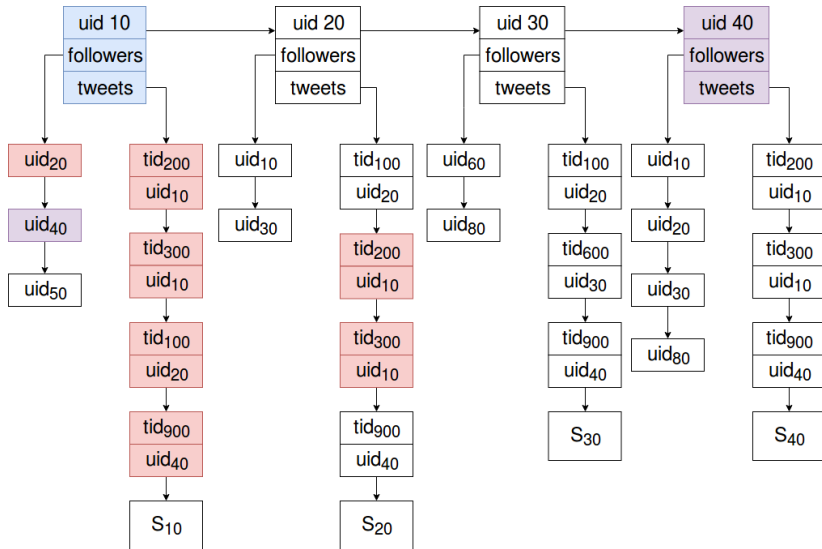


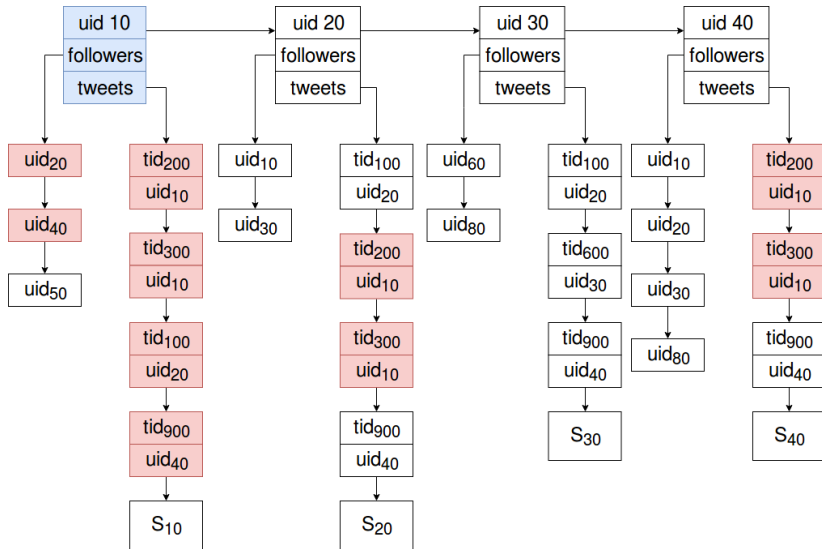


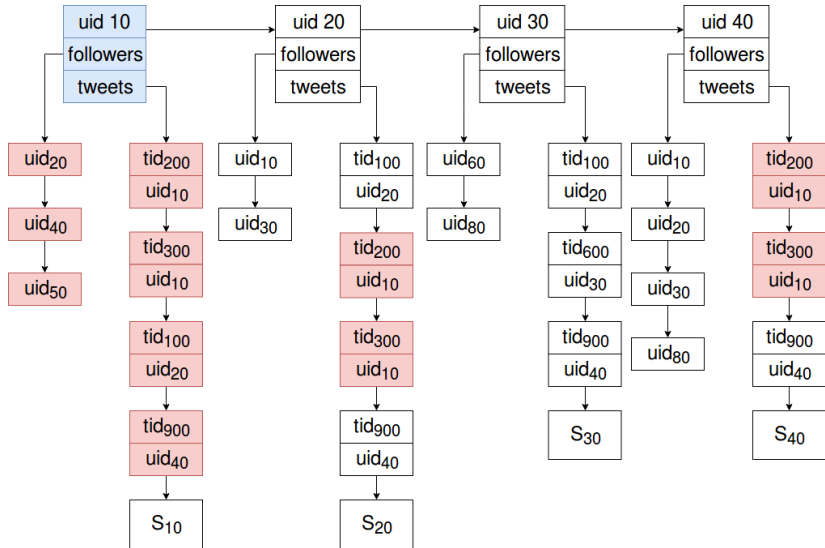


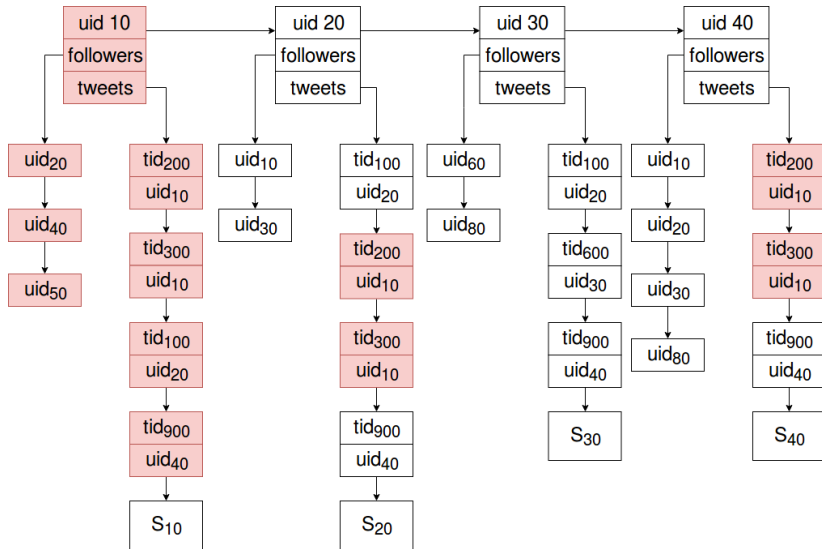












Lookup

Αναζήτηση ενός tweet με αναγνωριστικό *tid* στη λίστα των tweets του χρήστη με αναγνωριστικό *uid*.
Έχει την εξής μορφή:

L uid tid

Ορίσματα:

uid Αναγνωριστικό του χρήστη

tid Αναγνωριστικό του tweet

Lookup

Κατά το γεγονός αυτό:

1. Θα αναζητήσετε στη λίστα χρηστών το χρήστη με αναγνωριστικό *uid*
2. Θα αναζητήσετε στη λίστα των tweets του τον κόμβο με αναγνωριστικό *tid*

Merge

Συγχώνευση των tweets στον πίνακα Hashtags που ανήκουν στην κατηγορία *hashtag₁* ή *hashtag₂*.

Έχει την εξής μορφή:

$$M \text{ } hashtag_1 \text{ } hashtag_2$$

Ορίσματα:

hashtag₁ Κατηγορία 1

hashtag₂ Κατηγορία 2

Merge

Κατά το γεγονός αυτό:

1. Θα αναζητήσετε στον πίνακα Hashtags τις δύο λίστες tweets που αντιστοιχούν στα παραπάνω hashtags
2. Θα συνενώσετε τις λίστες αυτές με βάση το πεδίο *timestamp*
3. **Προσοχή:** Η διαδικασία αυτή θα πρέπει να ολοκληρώνεται σε χρονική πολυπλοκότητα $O(n_1 + n_2)$, όπου n_1 και n_2 είναι το πλήθος των κόμβων των λιστών tweets με hashtags $hashtag_1$ και $hashtag_2$, αντίστοιχα

Filter

Αναζήτηση των tweets που ανήκουν στην κατηγορία hashtag, τα οποία ήταν στην επικαιρότητα (trend) την ίδια χρονική περίοδο με το tweet με αναγνωριστικό *tid*.

Έχει την εξής μορφή:

F tid hashtag

Ορίσματα:

tid Αναγνωριστικό του tweet

hashtag Κατηγορία του tweet

Filter

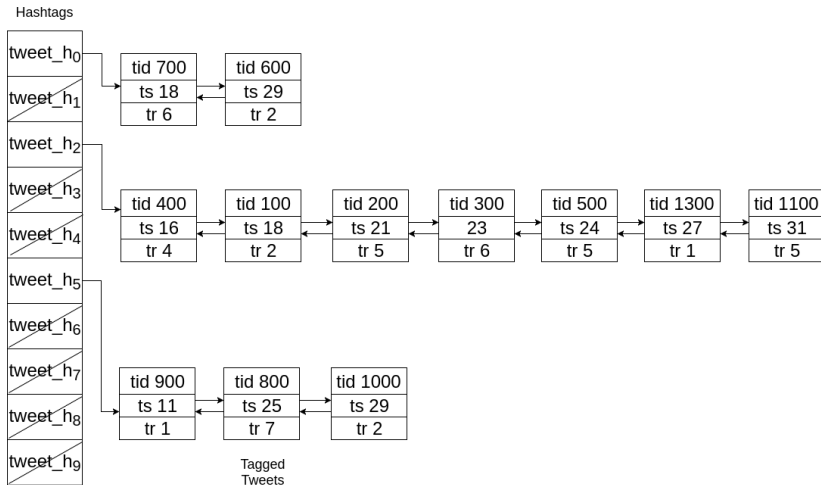
Κατά το γεγονός αυτό:

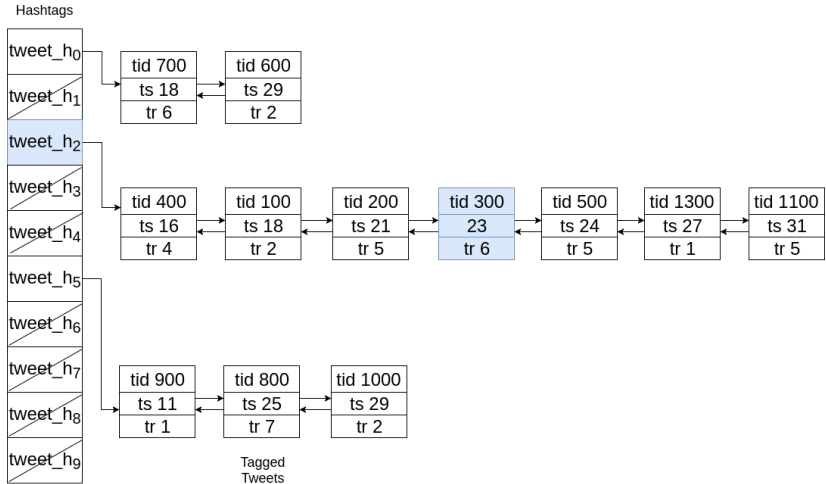
1. Θα αναζητήσετε στον πίνακα Hashtags τη λίστα των tweets της κατηγορίας hashtags
2. Θα εντοπίσετε το tweet με αναγνωριστικό *tid*
3. Θα εξετάσετε το πεδίο *time_relevance*
4. Ξεκινώντας από το tweet αυτό, θα κάνετε κατάλληλη διάσχιση zig-zag στη λίστα προκειμένου να τυπώσετε τα tweets των οποίων το πεδίο *timestamp* έχει τιμή στο διάστημα:

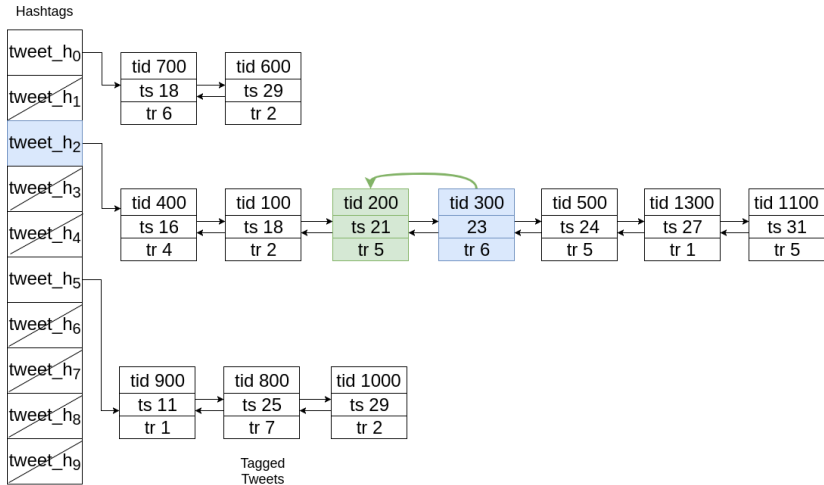
$$[timestamp_{tid} - time_rel_{tid}, timestamp_{tid} + time_rel_{tid}]$$

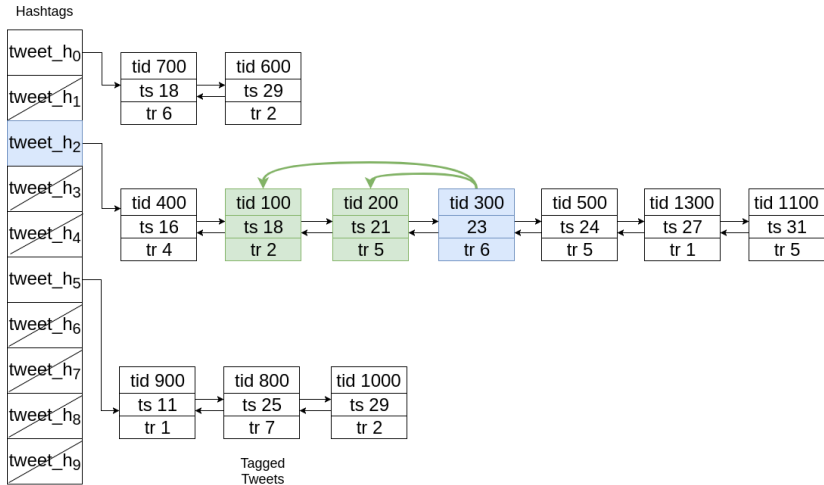
Filter: Παράδειγμα

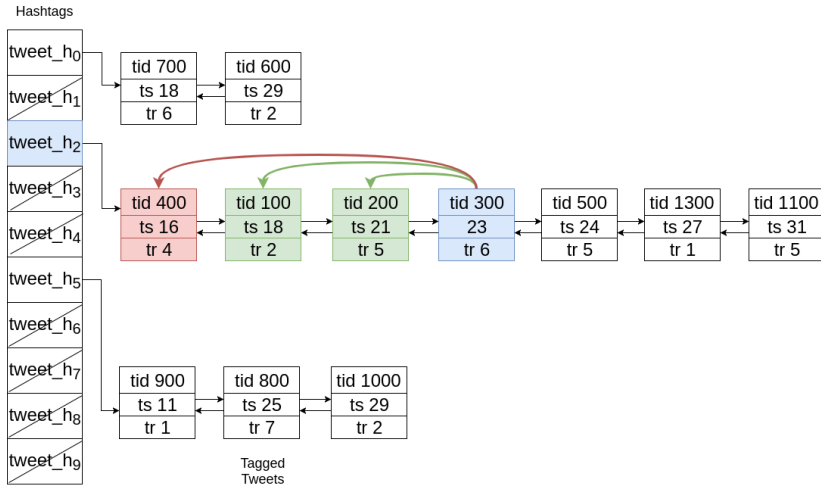
- ▶ Έστω γεγονός F 300 2

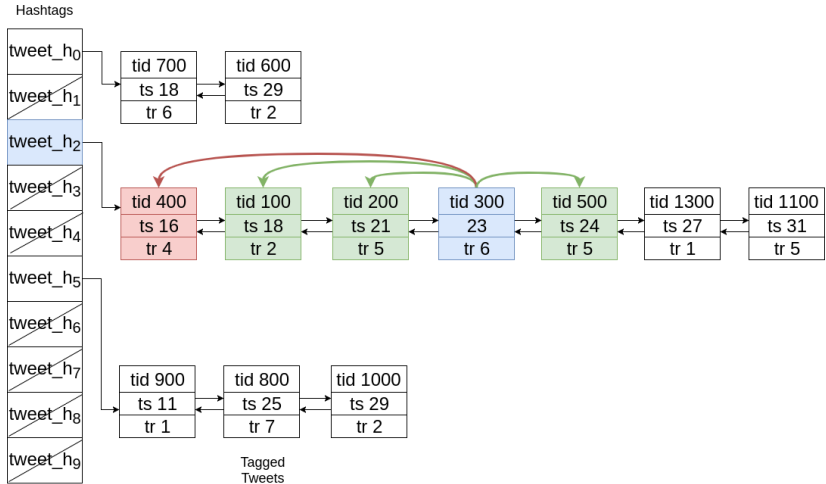




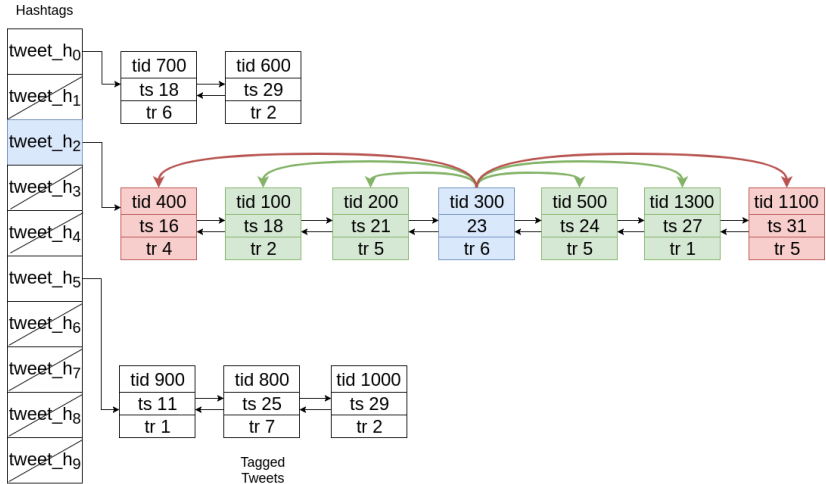












Print Users

Εκτύπωση όλων των χρηστών από τη λίστα χρηστών.
Έχει την εξής μορφή:

X

Print Tweets

Εκτύπωση όλων των tweets του πίνακα Hashtags.
Έχει την εξής μορφή:

Υ

Thanks!