

Διαβάθμιση επιστημόνων μέσω ανάλυσης σύνθετων δικτύων.

Complex network analysis for ranking Scientists

Διπλωματική εργασία
ΤΟΥ
Ελευθέριου Δ. Μπαρμπαγιάννη
Βόλος, Οκτώβρης 2016

Επιβλέποντες:

Δημήτριος Κατσαρός

Επίκουρος Καθηγητής Π.Θ.

Παναγιώτης Μποζάνης

Καθηγητής Π.Θ.

Εγκρίθηκε από την διμελή εξεταστική επιτροπή την ημερομηνία εξέτασης.

(Υπογραφή)

(Υπογραφή)

.....

.....

Δημήτριος Κατσαρός

Παναγιώτης Μποζάνης

Επίκουρος Καθηγητής Π.Θ.

Καθηγητής Π.Θ.

(Υπογραφή)

.....

Ελευθέριος Δ. Μπαρμπαγιάννης

Διπλωματούχος Μηχανικός Ηλεκτρονικών Υπολογιστών, Τηλεπικοινωνιών και Δικτύων
του Τμήματος Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών,

Πανεπιστημίου Θεσσαλίας

©2016 – All rights reserved

Ευχαριστίες

Μετά την περάτωση της εργασίας αυτής, θα ήθελα να ευχαριστήσω θερμά τους επιβλέποντες κ. Δημήτριο Κατσαρό και κ. Παναγιώτη Μποζάνη για την εμπιστοσύνη που μου έδειξαν, την άριστη συνεργασία και τις υποδείξεις που διευκόλυναν την εκπόνηση της διπλωματικής εργασίας.

Θα ήθελα να ευχαριστήσω του γονείς μου που όλα αυτά τα χρόνια με στήριξαν, με βοήθησαν και με εμπιστεύτηκαν. Την Χριστίνα που ήταν ένα πολύτιμο στήριγμα στα τελευταία χρόνια των σπουδών μου. Τους συντρόφους μου που ο καθένας από την πλευρά του με βοήθησαν στο να φτάσω στο τέλος των σπουδών μου. Στον πολύ καλό μου φίλο Μιχάλη που χάρη σε αυτόν κατάφερα να ξεπεράσω δυσκολίες κατά τη διάρκεια των σπουδών μου και αυτής της εργασίας. Σε όλους τους φίλους μου.

Ελευθέριος Μπαρμπαγιάννης,

Βόλος 2016

Περίληψη

Σκοπός της παρούσας εργασίας είναι η μελέτη δεικτών, που χρησιμοποιούνται για την μέτρηση της «αξίας» ενός επιστήμονα μέσω της επιρροής που ασκούν οι ερευνητικές τους εργασίες, για να προσπαθήσουμε να εισάγουμε temporal aspects στην έννοια των coterminal citation. Αυτό θα γίνει με την δημιουργία ενός νέου δείκτη που παίρνει ως βάση τον F-index και ονομάζεται trend F-index.

Abstract

Goal of this thesis is to study indecies, which use as metrics to measure the “level of excellence” of acientist, through the influence of their research papers, try to introduce temporal aspects in coterminal citation meaning. We introduce a new index, that index based on F-index and call it trend F-index.

Περιεχόμενα

1 Εισαγωγή

1.1 Θέμα

1.2 Σκοπός - Σημασία

1.3 Προβλήματα

2 Βάσεις δεδομένων επιστημονικών δημοσιεύσεων

3 Βιβλιογραφία στους επιστημονομετρικούς δείκτες

3.1 H-index

3.1.1 Contemporary H-index

3.1.2 Trend H-index

3.1.3 Normalized H-index

3.1.4 Yearly H-index

3.1.5 Normalized yearly H-index

3.2 F-index

4 Trend F-index

4.1 Στόχος

4.2 Βάση Δεδομένων

4.3 Γλώσσα Προγραμματισμού Python

4.4 Αλγόριθμος

5 Αποτελέσματα

6 Συζήτηση – Συμπεράσματα

Βιβλιογραφία

Εισαγωγή

1.1 Η αξιολόγηση του έργου ενός επιστήμονα έχει προσελκύσει μεγάλο ενδιαφέρον, εξαιτίας των πλεονεκτημάτων που δίνει ένας αμερόληπτος και δίκαιος δείκτης. Ένα τέτοιος δείκτης μπορεί να αξιοποιηθεί για την σύγκριση μεταξύ επιστημόνων, όσο αφορά την πρόσληψη διδακτικού προσωπικού σε πανεπιστήμια, ερευνητικά κέντρα και την βιομηχανία, την απόδοση ενός βραβείου, την χρηματοδότηση ερευνητικών προτάσεων κτλ..

Αποκτά ενδιαφέρον όσο αφορά την ποσοτικοποίηση της αξίας του ερευνητικού έργου που παράγουν, όχι μόνο ένας μεμονωμένος επιστήμονας, αλλά και ένας ολόκληρος ερευνητικός οργανισμός, μια ομάδα ερευνητών.

Παρά το γεγονός ότι, το θέμα της κατάταξης ενός επιστήμονα ή ενός περιοδικού χρονολογείται από το εβδομήντα με την πρωτοπόρα εργασία του Garfield και συνεχίστηκε με κάποιες δημοσιεύσεις, την τελευταία δεκαετία έχουμε γίνει μάρτυρες μιας «άνθησης» αυτού του πεδίου λόγω της διάδοσης των ψηφιακών βιβλιοθηκών, το οποίο έκανε διαθέσιμο ένα τεράστιο ποσό των βιβλιογραφικών δεδομένων.

Αντικείμενο της επιστημονομετρικής ανάλυσης είναι η καταγραφή και επεξεργασία δεδομένων που σχετίζονται με τις επιστημονικές δημοσιεύσεις και η εξαγωγή των σχετικών επιστημονομετρικών δεικτών, όπως ο αριθμός των δημοσιεύσεων, οι αναφορές σε αυτές από άλλες δημοσιεύσεις, ο συσχετισμός του γ με συγκεκριμένους φορείς, επιστημονικά πεδία, κλπ..

1.2 Μια διαδομένη μέθοδος αξιολόγησης της επιστημονικής εργασίας είναι ο ορισμός μιας αντικειμενικής συνάρτησης που υπολογίζει κάποια βαθμολογία για τα υπό αξιολόγηση αντικείμενα., λαμβάνοντας υπόψη το γράφημα που δημιουργείται από τις αναφορές (citations) του δημοσιευμένου άρθρου.

Φυσικά ο καθορισμός και η παραγωγή μια ποιοτικής και αντιπροσωπευτικής μέτρησης δεν είναι ένα εύκολο έργο, δεδομένου ότι θα πρέπει να παρθεί υπόψη το σύνολο του αντίκτυπου των εργασιών του επιστήμονα. Κατά κύριο λόγο οι περισσότερες μέθοδοι βασίζονται στην επεξεργασία του συνολικού αριθμού των δημοσιεύσεων, των αναφορών που έχουν λάβει αυτές οι δημοσιεύσεις, τον μέσο αριθμό των αναφορών ανά δημοσίευση, τον μέσο αριθμό των δημοσιεύσεων ανά έτος κτλ..

Κύριο πλεονέκτημα αυτών των μεθόδων είναι ότι μπορούν να εκμεταλλευτούν, με αυτοματοποιημένο τρόπο, όλο τον πλούτο των δημοσιεύσεων και των αναφορών που υπάρχουν στις ψηφιακές βιβλιοθήκες.

Με βάση τα στοιχεία των επιστημονικών δημοσιεύσεων είναι δυνατός ο προσδιορισμός χαρακτηριστικών και τάσεων της ερευνητικής παραγωγής σε επίπεδο φορέα, χώρας ή ευρύτερου συνόλου χωρών, η εκτίμηση της απήχησης του επιστημονικού έργου, η αξιολόγηση της ερευνητικής δραστηριότητας και ο εντοπισμός εθνικών και πολυεθνικών δικτύων μεταξύ επιστημόνων και επιστημονικών κλάδων.

1.3 Οι μέθοδοι που χρησιμοποιούνται μέχρι τώρα στην ανάλυση των δημοσιεύσεων και των αναφορών που έχουν, ακόμη και οι πιο δημοφιλείς όπως ο H-index και ο PageRank, παρουσιάζουν ένα ή περισσότερα μειονεκτήματα, όπως το να μην μετράνε τη σημασία ή τον αντίκτυπο των εγγράφων, να επηρεάζονται από ένα μικρό αριθμό δημοσιεύσεων που είχαν μεγάλη επιτυχία όσο αφορά τις αναφορές και οι υπόλοιπες δημοσιεύσεις να μην έχουν τέτοιο αντίκτυπο, να μην μετρούν την παραγωγικότητα, να επηρεάζονται από τις «αυτοαναφορές» που γίνονται σε δημοσιεύσεις.

Επίσης ρόλο παίζει το επιστημονικό πεδίο που αφορά η δημοσίευση π.χ. στον τομέα της πληροφορικής υπάρχει υψηλή παραγωγικότητα, ενώ στον τομέα των κοινωνικών επιστημών οι ρυθμοί δημοσίευσης είναι πολύ πιο αργοί.

Σημαντικό ρόλο παίζει το χρονικό διάστημα που έχει γίνει η δημοσίευση για τις αναφορές που έχει λάβει. Ο αριθμός των αναφορών σε μια επιστημονική εργασία εξαρτάται από το χρονικό διάστημα που έχει παρέλθει μετά τη δημοσίευση της. Συνήθως οι παλαιότερες δημοσιεύσεις έχουν και περισσότερες αναφορές, χωρίς αυτό να συνδέεται πάντα με την απήχυσή τους στην επιστημονική κοινότητα.

Γενικότερα η χρήση τέτοιων δεικτών είναι αμφιλεγόμενη, δεδομένου ότι μια τέτοιου είδους εκτίμηση αποτελεί μια σύνθετη επιστημονική και κοινωνική διαδικασία, η οποία δεν μπορεί να περιοριστεί σε ένα μόνο δείκτη επιστημονικών μετρήσεων. Αυτός όμως ο προβληματισμός δεν αναιρεί τη σημασία των επιστημονομετρικών δεικτών ως πολύτιμη πηγή δεδομένων και, όπως άλλωστε ισχύει με την ερμηνεία των περισσότερων δεικτών, μπορούν να ξεπεραστούν όταν οι επιστημονομετρικοί δείκτες ερμηνευτούν στο σωστό πλαίσιο.

ΒΑΣΕΙΣ ΔΕΔΟΜΕΝΩΝ

ΕΠΙΣΤΗΜΟΝΙΚΩΝ ΔΗΜΟΣΙΕΥΣΕΩΝ

Διεθνώς, οι πλέον καθιερωμένες βάσεις δεδομένων που περιλαμβάνουν βιβλιογραφικές εγγραφές επιστημονικών δημοσιεύσεων σε παγκόσμιο επίπεδο και στοιχεία για τις αναφορές μεταξύ τους, είναι τα συστήματα Web of Science (της εταιρίας Thomson Reuters), Scopus (της Elsevier) και Google Scholar (της Google).

Στο σύστημα Web of Science (WoS) ευρετηριάζονται πάνω από 12.000 περιοδικά τα οποία υπόκεινται σε αξιολόγηση κριτών ενώ από το 1990 προστέθηκε και η καταγραφή ορισμένων πρακτικών συνεδρίων. Είναι η παλαιότερη βάση δεδομένων επιστημονικών δημοσιεύσεων με βιβλιογραφικές εγγραφές και αναφορές που χρονολογούνται, για ορισμένους επιστημονικούς κλάδους, από το 1900. Περιλαμβάνονται αναλυτικά μεταδεδομένα για άρθρα, συγγραφείς και ερευνητικούς οργανισμούς. Σημαντικό πλεονέκτημα του Web of Science θεωρείται η αξιοπιστία του, αποτέλεσμα της αυστηρής αξιολόγησης των εκδόσεων και κυρίως των περιοδικών που εισάγονται στο σύστημα βάσει συγκεκριμένων κριτηρίων, μεταξύ των οποίων και η επιστημονική τους απήχηση.

Στο σύστημα Scopus ευρετηριάζονται 18.500 τίτλοι περιοδικών, πρακτικά συνεδρίων και βιβλία. Όπως και το σύστημα WoS, τα μεταδεδομένα περιέχουν αναλυτικά στοιχεία για άρθρα, συγγραφείς και οργανισμούς, ενώ η εισαγωγή των εκδόσεων πραγματοποιείται μετά από ποιοτική αξιολόγηση βάσει κριτηρίων. Το σύστημα Scopus παρέχει, σε σχέση με το σύστημα WoS, πιο ισορροπημένη γεωγραφική κάλυψη, υστερεί όμως χρονικά, οι παλαιότερες βιβλιογραφικές εγγραφές ξεκινούν το 1966, ενώ οι αναφορές σε δημοσιεύσεις καταγράφονται μετά το 1995.

Το σύστημα Google Scholar περιλαμβάνει ένα τεράστιο αριθμό πηγών, πολλές από τις οποίες αφορούν ακαδημαϊκές εργασίες, διατριβές, τεχνικές εκθέσεις, παραδοτέα

ερευνητικών έργων κτλ.. Παρά το μεγάλο εύρος σε περιεχόμενο, τα μεταδεδομένα των εγγραφών είναι περιορισμένα και απουσιάζουν κρίσιμα στοιχεία για τη γεωγραφική ή τη θεματική κάλυψη που παρέχει το σύστημα.

Με βάση τα παραπάνω χαρακτηριστικά, το σύστημα Google Scholar, παρά τον τεράστιο αριθμό πηγών που περιλαμβάνει, δεν είναι κατάλληλο για βιβλιομετρικές αναλύσεις που αναφέρονται σε επίπεδο χωρών ή οργανισμών, λόγω της έλλειψης μεταδεδομένων που απαιτούνται για την ταυτοποίηση των δημοσιεύσεων και της απουσίας κριτηρίων που διασφαλίζουν την ποιότητα των δημοσιεύσεων που περιλαμβάνονται στο σύστημα.

Όσον αφορά τα συστήματα Web of Science και Scopus, και τα δυο διασφαλίζουν τη διάθεση αναλυτικών μεταδεδομένων και την ποιότητα των δημοσιεύσεων που περιλαμβάνουν και είναι αυτά που χρησιμοποιούνται διεθνώς για την υλοποίηση βιβλιομετρικών μελετών. Οι βάσεις δεδομένων των δυο αυτών συστημάτων επικαιροποιούνται σε συνεχή βάση, διευρύνουν τον αριθμό και το εύρος των επιστημονικών περιοδικών που ευρετηριάζουν ενώ εμπλουτίζουν τις πηγές δεδομένων συμπεριλαμβάνοντας πρακτικά συνεδρίων, βιβλία, μονογραφίες κλπ.. Ωστόσο και οι δυο βάσεις δεν επιτυγχάνουν να παρέχουν πλήρη κάλυψη των επιστημονικών δημοσιεύσεων, τόσο γεωγραφικά (αναφέρεται ότι η WoS έχει καλύτερη κάλυψη των αγγλόφωνων περιοδικών και κυρίως των ΗΠΑ, ενώ η Scopus έχει καλύτερη κάλυψη των ευρωπαϊκών περιοδικών) όσο και θεματικά στα διάφορα επιστημονικά πεδία (αναφέρεται ότι η WoS υπερτερεί στην κάλυψη των επιστημονικών περιοχών Natural Sciences ενώ η Scopus των Medical Sciences).

ΒΙΒΛΙΟΓΡΑΦΙΑ ΣΤΟΥΣ

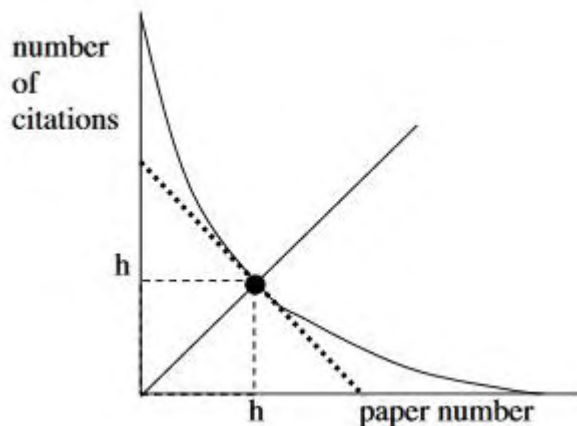
ΕΠΙΣΤΗΜΟΝΟΜΕΤΡΙΚΟΥΣ ΔΕΙΚΤΕΣ

3.1 Για να ξεπεραστούν συλλογικά όλα αυτά τα μειονεκτήματα των μετρήσεων, το 2005 ο J.E. Hirsch πρότεινε τον πρωτοποριακό H-index, ο οποίος σε σύντομο χρονικό διάστημα, έγινε εξαιρετικά δημοφιλής.

Ο ορισμός του H-index είναι ο εξής:

«Ένας συγγραφέας εμφανίζει δείκτη h όταν h από τις N_p εργασίες του έχουν τουλάχιστον h αναφορές η καθεμία, ενώ οι υπόλοιπες (N_p-h) εργασίες του έχουν λιγότερες από h αναφορές η καθεμία»

Στην εικόνα παρατηρούμε, ότι η τομή της γραμμής 45 μοιρών με την καμπύλη, που προκύπτει από τον αριθμό των αναφορών (number of citation) προς τον αριθμό των δημοσιεύσεων (paper number) μας δίνει τον αριθμό h .



Για παράδειγμα, αν μια συγκεκριμένη ομάδα εργασιών ή συγκεκριμένες εργασίες κάποιου συγγραφέα εμφανίζουν $h=12$, αυτό σημαίνει ότι σε αυτή την ομάδα των εργασιών υπάρχουν 12 εργασίες που έχουν τουλάχιστον 12 (ή και περισσότερες)

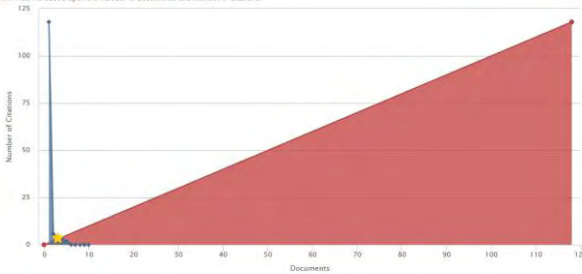
αναφορές η καθεμία. Δηλαδή, ο h-index εξάγεται πρακτικά με κατάταξη της συγκεκριμένης ομάδας των εργασιών κατά φθίνουσα σειρά ξεκινώντας από αυτή που έχει τις περισσότερες αναφορές. Καθώς προχωρούμε προς τα κάτω, στο σημείο που ο αριθμός των αναφορών που έχει λάβει κάποια εργασία γίνει μικρότερος από τον αύξοντα αριθμό της εργασίας, τραβάμε μία γραμμή και λαμβάνουμε ως h-index τον αμέσως προηγούμενο αύξοντα αριθμό. Εξ' ορισμού η τιμή του h-index εξαρτάται από τη βάση δεδομένων στην οποία θα γίνει η αναζήτηση.

Αυτή η μέτρηση υπολογίζει το εύρος της ερευνητικής εργασίας ενός επιστήμονα. Ο δείκτης αντιπροσωπεύει τόσο την παραγωγικότητα όσο και τον αντίκτυπο. Για ένα ερευνητή, το να έχει μεγάλο h-index, σημαίνει ότι έχει πολλά «καλά» άρθρα, και όχι μόνο μερικά «καλά» άρθρα.

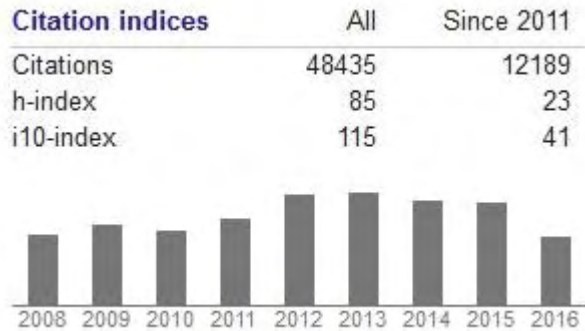
Η ανακάλυψη του h-index ήταν μια σημαντική ανακάλυψη στην ανάλυση παραπομπών. Αν και πολλές πτυχές της αναποτελεσματικότητας του αρχικού h-index είναι εμφανείς, όπως το να αποδίδει την ίδια σημασία σε όλες τις αναφορές και να μην δίνει σημασία στην ηλικία των αναφορών. Επίσης, ο h-index αποδίδει την ίδια σημασία σε όλες τα άρθρα, καθιστώντας έτσι τους νέους ερευνητές να έχουν ένα σχετικά μικρό h-index, επειδή δεν έχουν πολλά χρόνια που κάνουν δημοσιεύσεις ή το χρόνο να συσσωρεύσουν μεγάλο αριθμό παραπομπών για τα καλά τους χαρτιά. Έτσι ο h-index δεν μπορεί να αποκαλύψει λαμπρούς νέους επιστήμονες.

Στις παρακάτω εικόνες βλέπουμε πως εμφανίζεται ο h-index στο προφίλ ενός επιστήμονα (Alan Turing), στο Scopus στην 1^η εικόνα και στο Google Scholar στην 2^η εικόνα.

This author's h-index is 3
The h-index is based upon the number of documents and number of citations.



Note: Scopus is in progress of updating pre-1996 cited references going back to 1970. The h-index might increase over time.



3.1.1 Contemporary h-index. Είναι ικανός να αποκαλύψει λαμπρούς νέους επιστήμονες. Ο αρχικός h-index δεν λαμβάνει υπόψη την ηλικία ενός αντικειμένου. Μια τέτοια περίπτωση είναι ένας επιστήμονας που συνέβαλε σε μια σειρά σημαντικά άρθρα που παράγουν μεγάλο h-index, αλλά τώρα είναι ανενεργός ή συνταξιούχος. Ως εκ τούτου, ανώτεροι επιστήμονες, οι οποίοι εξακολουθούν να συμβάλλουν στις μέρες μας, ή νέοι λαμπροί επιστήμονες, οι οποίοι αναμένεται να συνεισφέρουν ένα μεγάλο αριθμό σημαντικών έργων στο άμεσο μέλλον, αλλά τώρα έχουν μόνο ένα μικρό αριθμό σημαντικών άρθρων, λόγω του μικρού διαστήματος, δεν διακρίνονται από τον αρχικό h-index. Έτσι προέκυψε η ανάγκη καθορισμού μιας γενίκευσης του h-index για τέτοιους είδους περιπτώσεις.

Ορίζεται ένα νέο σκορ για ένα άρθρο i με βάση τις αναφορές που έχει λάβει:

$$S^c(i) = \gamma * (Y(now) - Y(i+1))^{-\delta} * |C(i)|,$$

όπου $Y(i)$ είναι η χρονιά δημοσίευσης του άρθρου και $C(i)$ είναι τα άρθρα που αναφέρονται στο άρθρο i . Με $\delta = 1$, το $S^c(i)$ αποτελεί τον αριθμό των αναφορών που το άρθρο i έχει δεχθεί, διαιρεμένο με την ηλικία του. Ο συντελεστής γ μειώνεται με την παλαιότητα του άρθρου (π.χ. για $\gamma=4$, ένα άρθρο αυτής της χρονιάς έχει τιμή 4, ένα άρθρο 4 χρόνια πριν τιμή 1, ένα άρθρο 6 χρόνια πριν $4/6$ κτλ.)

Έτσι προκύπτει ένα παλιό άρθρο να χάνει την αξία του, παρά τις αναφορές που του έχουν γίνει και έτσι παίρνονται υπόψη περισσότερο τα νεότερα άρθρα. Έτσι προκύπτει ο εξής ορισμός για τον contemporary h-index:

«Ένας ερευνητής εμφανίζει contemporary h-index h^c , όταν h^c από τις N_p εργασίες του έχουν σκορ $S^c(i) \geq h^c$ η καθεμία, ενώ οι υπόλοιπες ($N_p - h^c$) εργασίες του έχουν σκορ $S^c(i) \leq h^c$. »

3.1.2 Trend h-index. Αποκαλύπτει τους επιστήμονες που το έργο τους αυτή την περίοδο μελετάται και αξιοποιείται. Το αρχικό h-index δεν λαμβάνει υπόψη το χρόνο, που ένα άρθρο απέκτησε μια συγκεκριμένη αναφορά. Για παράδειγμα ένας ερευνητής που συνέβαλε στην ερευνητική κοινότητα με μια σειρά από σημαντικά άρθρα κατά της διάρκειας της δεκαετίας του 1960 και πήρε πολλές αναφορές εκείνη την δεκαετία, θα έχει υψηλό h-index εξαιτίας των έργων που έχει κάνει στο παρελθόν. Εάν αυτά τα άρθρα δεν αναφέρονται πια είναι ένδειξη ενός ξεπερασμένου θέματος ή μιας ξεπερασμένης λύσης. Από την άλλη πλευρά αν αυτά τα άρθρα συνεχίζουν να αναφέρονται, τότε έχουμε την περίπτωση ενός σπουδαίου μυαλού που συνεχίζει να συμβάλει στην επιστημονική κοινότητα. Ταυτόχρονα μέσα από την γήρανση των παραπομπών, γεννιέται η δυνατότητα της αποκάλυψης πρωτοποριακών και καινοτόμων επιστημόνων που το έργο τους είναι «trendy», έτσι τα έργα τους αναφέρονται πολύ συχνά. Σε αντίθεση με τον contemporary h-index, εδώ ασχολούμαστε με την ηλικία των αναφορών των άρθρων και όχι με την ηλικία των άρθρων. Με αυτό τον τρόπο γίνεται εκτίμηση του αντίκτυπου του άρθρου σε μια συγκεκριμένη χρονική στιγμή. Έτσι ορίζεται η παρακάτω εξίσωση:

$$S^t(i) = \gamma \sum (Y(\text{now}) - Y(x) + 1)^{\delta} \quad (\text{για όλα τα } x \text{ που ανήκουν στο } C(i)),$$

όπου γ , δ , $Y(i)$ και $S(i)$ για κάθε άρθρο ορίζονται όπως πριν. Έτσι προκύπτει ο εξής ορισμός για τον trend h-index:

«Ένας ερευνητής εμφανίζει trend h-index h^t , όταν h^t από τις N_p εργασίες του έχουν σκορ $S^t(i) \geq h^t$ η καθεμία, ενώ οι υπόλοιπες ($N_p - h^t$) εργασίες του έχουν σκορ $S^t(i) \leq h^t$: »

3.1.3 Normalized h-index. Αποτελεί μια κανονικοποιημένη μορφή του h-index, μιας και οι επιστήμονες δεν δημοσιεύουν τον ίδιο αριθμό άρθρων. Για αυτό έχουμε τον παρακάτω ορισμό:

«Ένας ερευνητής εμφανίζει normalized h-index $h^n = h/N^p$, όταν h από τις N_p εργασίες του έχουν τουλάχιστον h αναφορές η καθεμία, ενώ οι υπόλοιπες ($N_p - h$) εργασίες του δεν έχουν λάβει παραπάνω από h αναφορές»

3.1.4 Yearly h-index. Αποτελεί κατάλληλο δείκτη για την κατάταξη περιοδικών, συνεδρίων και έχει τον εξής ορισμό:

«Ένα συνέδριο ή ένα περιοδικό εμφανίζει yearly h-index h_y για τη χρονιά y αν h_y από τα άρθρα $N_{p,y}$ που έχουν δημοσιευτεί κατά τη διάρκεια της χρονιάς y έχουν λάβει τουλάχιστον h_y αναφορές το καθένα, και τα υπόλοιπα ($N_{p,y} - h_y$) άρθρα έχουν λάβει όχι περισσότερες από h_y αναφορές».

3.1.5 Normalized yearly h-index. Αποτελεί μια κανονικοποιημένη μορφή του yearly h-index, για να αντισταθμιστεί το πρόβλημα ότι τα συνέδρια δεν δημοσιεύουν ακριβώς τον ίδιο αριθμό από άρθρα. Ο ορισμός του δείκτη είναι ο εξής:

«Ένα συνέδριο ή ένα περιοδικό για τη χρονιά y εμφανίζει normalized yearly h-index $h_y^n = h_y/N_{p,y}$ αν h_y από τα άρθρα $N_{p,y}$ που έχουν δημοσιευτεί κατά τη διάρκεια της χρονιάς y έχουν λάβει τουλάχιστον h_y αναφορές το καθένα, και τα υπόλοιπα ($N_{p,y} - h_y$) άρθρα έχουν λάβει όχι περισσότερες από h_y αναφορές».

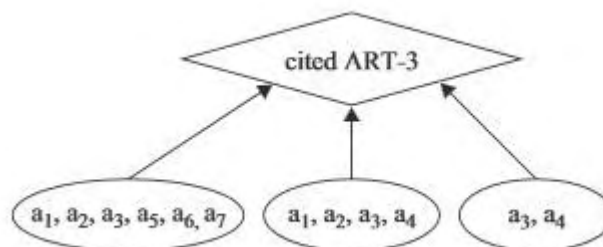
3.2 Σε όλους τους δείκτες, που έχουν αναπτυχθεί μέχρι σήμερα, υπάρχει μια αδυναμία, είτε αφορά την κατάταξη ατόμων είτε την κατάταξη συνεδρίων. Ο h-index

αποτελεί ένα θύμα αυτής της επιπλοκής. Η ανεπάρκεια των δεικτών πηγάζει από τις coterminal citations, δηλαδή αυτόαναφορές που αποτελούν ένα αμφιλεγόμενο κομμάτι όσο αφορά τον πραγματικό αντίκτυπο ενός άρθρου.

Λύση σε αυτό το πρόβλημα έρχεται να δώσει ο f-index, που δεν προσπαθεί να αποκλείσει τις αυτοαναφορές, αλλά μέσα από την έννοια των coterminal citations γίνεται προσπάθεια ανακάλυψης των μοντέλων και ποσοτικοποίησης αυτής της τιμής.

Ουσιαστικά ομαδοποιούνται οι επιστήμονες που με εργασίες αναφέρονται στο άρθρο, ανάλογα με το πόσες φορές έχουν αναφερθεί στο συγκεκριμένο άρθρο, έτσι δημιουργούνται σετ τα οποία περιλαμβάνουν το πλήθος των συγγραφέων που αναφέρθηκαν i φορές σε αυτό το άρθρο. Από αυτά τα σετ δημιουργείται ένα διάνυσμα f^A , όπου A το άρθρο που αναφέρονται οι συγγραφείς, που είναι ίσο με $f^A = \{f_1^A, f_2^A, f_3^A, \dots, f_{nca}^A\}$ (nca : αριθμός άρθρων που αναφέρονται στο άρθρο A), από $f_i^A =$ πλήθος συγγραφέων που αναφέρθηκαν i φορές στο άρθρο/ συνολικό αριθμός όλων των διαφορετικών συγγραφέων που αναφέρονται στο άρθρο. Χρησιμοποιώντας το διάνυσμα $s_1 = \{nca, nca-1, \dots, 1\}$, παίρνουμε τον αριθμό $N_f^A = f^A s_1$. Οπότε προκύπτει ο ορισμός του f-index, με ανάλογο τρόπο όπως ο h-index:

«Για να υπολογιστεί ο f-index ενός ερευνητή, υπολογίζονται οι ποσότητες $N_f^{A_i}$ των άρθρων του A_i και τα ταξινομούνται σε φθίνουσα σειρά. Το σημείο που η θέση στη σειρά γίνεται μεγαλύτερη από το αντίστοιχο $N_f^{A_i}$ στην ταξινομημένη σειρά, ορίζει τον f-index για αυτό τον ερευνητή.»



Για παράδειγμα το άρθρο της εικόνας αποκτά αυτά τα σετ: $F^A_1=\{a5,a6,a7\}$, $F^A_2=\{a1,a2,a4\}$, $F^A_3=\{a3\}$. Έτσι προκύπτει το διάνυσμα $f^{ART-3}=\{3/7,3/7,1/7\}$ και ο τελικός αριθμός του άρθρου $N^A_f=f^A \cdot s_1=3/7 \cdot 3+3/7 \cdot 2+1/7 \cdot 1=167 \approx 2.28$. Οπότε δουλεύοντας με την ίδια μέθοδο για όλα τα άρθρα και εφαρμόζοντας τον ορισμό του f-index προκύπτει ο f-index του ερευνητή.

Trend F-index

4.1 Στόχος της εργασίας είναι να εισαχθούν temporal aspects στην έννοια των coterminal citation για την αποφυγή πληθωριστικών τάσεων.

Για αυτό το λόγο αξιοποιήθηκαν πλευρές των εργασιών του f-index και του trend h-index, ώστε να καταφέρουμε να αποκαλύψουμε επιστήμονες που το έργο τους μελετάται και αξιοποιείται σε μια χρονική στιγμή, αντιμετωπίζοντας αυτό το μειονέκτημα του trend h-index.

4.2 Ως βάση δεδομένων χρησιμοποιήθηκε ένα αρχείο με 251 επιστήμονες, που περιλάμβανε το όνομα του κάθε επιστήμονα, τις δημοσιεύσεις του μαζί με την χρονιά που έγιναν και τις αναφορές που είχαν οι δημοσιεύσεις, μαζί με τα ονόματα των επιστημόνων που τις έκαναν και την χρονιά που έγιναν.

4.3 Η γλώσσα προγραμματισμού που χρησιμοποιήθηκε είναι η Python, που αποτελεί ένα πολύ εύχρηστο εργαλείο και δυνατό εργαλείο για την ανάπτυξη κώδικα. Ενώ παρέχει και αρκετές λειτουργίες και βιβλιοθήκες όσο αφορά τις μαθηματικές προσεγγίσεις που έπρεπε να κάνουμε στον κώδικα.

4.4 Όσο αφορά την λειτουργία του αλγορίθμου, που φτιάχτηκε, ακολουθήθηκαν τα παρακάτω βήματα,

για κάθε δημοσίευση A, παίρνουμε μια- μια τις χρονιές (αρχική χρονιά που έχουν γίνει οι δημοσιεύσεις - τωρινή χρονιά που γίνεται η έρευνα) και συγκεντρώνουμε τους συγγραφείς όλων των αναφορών, αυτής της χρονιάς.

```

number_of_publ = data_sc[0]['number_of_publication']
cit_year = [year for year in range(STARTING_YEAR, PRESENT_YEAR + 1)] # initialize to search between specifically years
temp_pub = []
#calculate the index for every publication
for pub in range(number_of_publ):
    number_of_cit = data_sc[0]['publications'][pub]['number_of_citation']
    pub_index = 0
    for year in cit_year:
        temp_year = []
        nca_year = 0
        for cit in range(number_of_cit):
            if (data_sc[0]['publications'][pub]['citations'][cit]['year_of_citation']) == year:
                temp_year.append(data_sc[0]['publications'][pub]['citations'][cit]['citation_authors'])

```

Κατατάσσουμε τους συγγραφείς αυτή της χρονιάς σε ομάδες ανάλογα με τις αναφορές που έχουν κάνει στο συγκεκριμένο άρθρο αυτή την χρονιά F_i^{A-year} , όπου i το πλήθος των αναφορών που έχουν κάνει οι συγγραφείς που ανήκουν στο σύνολο, A η δημοσίευση που εξετάζουμε και $year$ η χρονιά που εξετάζουμε.

Μετά ορίζουμε την ποσότητα f_i^{A-year} , όπως στον f -index, που έχει τη μορφή $f_i^{A-year} = |F_i^{A-year}| / \text{συνολικό αριθμό των διαφορετικών συγγραφέων που έχουν κάνει αναφορά αυτή την χρονιά}$. Έτσι δημιουργείται το διάνυσμα $f^{A-year} = \{f_1^{A-year}, f_2^{A-year}, f_3^{A-year}, \dots, f_{nca}^{A-year}\}$, όπου nca ο αριθμός των αναφορών στο άρθρο την χρονιά $year$.

Αξιοποιούμε το διάνυσμα $s1 = \{nca, nca-1, \dots, 1\}$, που χρησιμοποιείται στον f -index, για να δημιουργήσουμε, αντίστοιχα βάρη σε κάθε διάνυσμα αναφορών, δίνοντας βάρος αντίστροφα με το πλήθος των αναφορών, δηλαδή μεγαλύτερο βάρος στην 1 αναφορά κτλ.. Έτσι δημιουργείται ο $N_f^{A-year} = f^{A-year} \cdot s1 = f_1^{A-year} * nca + f_2^{A-year} * (nca-1) + f_{nca}^{A-year} * 1$.

```
temp_year = [item for sublist_1 in temp_year for item in sublist_1]
y = Counter(temp_year)
distinct_authors_year = len(y)
#find the Nf for every year
nf_year = 0
for k in range(max(y.values())):
    nf_year += (sum(1 for x in y.values() if (x == k + 1)) / distinct_authors_year) * (nca_year - k)
```

Για να μπορέσουμε να δώσουμε διαφορετικό βάρος στην κάθε χρονιά, θέλοντας να ενισχύσουμε τις χρονιές που είναι πιο κοντά στην δική μας, αξιοποιούμε την μεταβλητή γ , που χρησιμοποιείται στον cotemporary h-index και στον trend h-index. Έτσι για κάθε χρονιά η μεταβλητή γ παίρνει την τιμή $\gamma = \text{αρχικό } \gamma / (\text{τωρινή χρονιά} - \text{την χρονιά που εξετάζουμε} + 1)$. Έτσι για την τωρινή χρονιά η μεταβλητή θα είναι ίση με γ , για την προηγούμενη χρονιά $\gamma = \gamma/2$, για την μεθεπόμενη $\gamma = \gamma/3$ κτλ.

```
#calculate the coeff for every year
coeff = GAMMA / (PRESENT_YEAR - year + 1)
#calculate the index for this year publication
pub_index += nf_year * coeff
```

Έτσι μπορούμε να έχουμε για κάθε δημοσίευση ένα index, που θα το αποθηκεύσουμε σε ένα πίνακα, θα τον ταξινομήσουμε σε φθίνουσα σειρά και θα έχουμε τον ορισμό του trend f-index, ανάλογο με αυτό του f-index.

```

        pub_index += nf_year * coeff
    temp_pub.append(round(pub_index, 2))
#sorted the publications in nonincreasing order
sorted_author_pubs = sorted(temp_pub, reverse=True)
#calculate trend_Findex
trend_F_index = calculate_trend_F_index(sorted_author_pubs, number_of_publ)

```

«Για να υπολογιστεί ο trend f-index ενός ερευνητή, υπολογίζονται οι ποσότητες N_f^{A-year} των άρθρων του A για την χρονιά year και τα ταξινομούνται σε φθίνουσα σειρά. Το σημείο που η θέση στη σειρά γίνεται μεγαλύτερη από το αντίστοιχο N_f^{A-year} στην ταξινομημένη σειρά, ορίζει τον trend f-index για αυτό τον ερευνητή.»

```

#calculate trend_findex like findex way
def calculate_trend_F_index(pub_data, pub_num):
    for i in range(1, pub_num+1):
        if (i > pub_data[i-1]):
            break
    return i-1

```

Για παράδειγμα, έχουμε το άρθρο A, που το αναφέρουν τα εξής άρθρα:

(α₁, α₂, α₃, α₅, α₆, α₇)(2008), (α₁, α₂, α₃, α₄)(2008)
 άρθρο)

(α₁, α₂, α₃, α₅, α₆, α₇)(2008), (α₁, α₂, α₃, α₄)(2008)

(α₁, α₂, α₃, α₄)(2010), (α₃, α₄)(2010)

Άρα παίρνουμε:

$$F_1^{A-2008} = [a4, a5, a6, a7], F_2^{A-2008} = [a1, a2, a3, a4]$$

$$F_1^{A-2010} = [a1, a2], F_2^{A-2010} = [a3, a4]$$

Οπότε έχουμε:

$$f^{A-2008} = \{4/7, 3/7\}, f^{A-2010} = \{2/4, 2/4\}$$

$$\text{και } N_f^{A-2008} = 4/7*2 + 3/7*1 = 1.14 + 0.42 = 1.56, N_f^{A-2010} = 2/4*2 + 2/4*1 = 1 + 0.5 = 1.5$$

Ως τελικό βήμα παίρνουμε, για $\gamma = 4$:

$$\text{Pub_index} = 4 * 1.56 + 4/3 * 1.5 = 8.24.$$

ΑΠΟΤΕΛΕΣΜΑΤΑ

Για τις ανάγκες της μέτρησης, βάλαμε περιορισμό, ως προς την χρονιά που θα αποτελούσε την χρονιά έναρξης, που ήταν το 1980 και αυτό γιατί υπήρχαν αναφορές με έτος δημοσίευσης το 0, που έτσι και αλλιώς θα ήταν άχρηστες για τον συγκεκριμένο αλγόριθμο και κατά 2^{ov} γιατί χρονολογίες πριν το 1980 δεν θα έδειχναν κάποια διαφορά στον αλγόριθμο. Ως «τωρινή» χρονιά για τον αλγόριθμο θεωρούμε το έτος 2010, μιας και η βάση δεδομένων είναι από εκείνη τη χρονιά. Το γ το πήραμε ίσο με 4, αξιοποιώντας την τιμή που δόθηκε στα πειράματα για τον trend h-index.

Παρατηρούμε στους πίνακες παρακάτω, ότι σε σχέση με τον F-index υπάρχουν ερευνητές που έχουν κερδίσει πολλές θέσεις. Για παράδειγμα ο Johnson που βρίσκεται στην 1^η θέση, από την 13^η θέση και ο Won Kim που βρίσκεται στην 2^η θέση από την 20^η θέση, σε σχέση με την κατάταξη τους στο πείραμα με τον F-index. Και οι 2 είναι ερευνητές που είχαν αρκετές αναφορές και το 2010 και το 2009. Ταυτόχρονα έχουμε αλλαγές και προς τα κάτω, όπως ο Tarjan που είναι στην 22^η θέση από την 3^η θέση και η Widom που είναι στην 16^η θέση από την 5^η θέση, μιας και το 2010 και το 2009 είχαν ελάχιστες αναφορές. Άλλο ένα τέτοιο παράδειγμα είναι ο DeWitt που είναι στη θέση 26, από την θέση 10 που ήταν.

Παρατηρούμε ακόμα ότι οι Shenker και Garcia-Molina, βρίσκονται στην πρώτη τετράδα, όπως και στον F-index, πράγμα που δείχνει ότι η δουλειά τους συνεχίζεται να δέχεται αρκετές αναφορές την περίοδο που μελετάμε.

1	David S. Johnson	108	23	Mario Gerla	64
2	Won Kim	94	23	David Maier	64
3	Hector Garcia-Molina	92	24	Wil van der Aalst	63
4	Scott Shenker	88	24	Martin Vetterli	63
4	Michael J. Black	88	24	Rakesh Agrawal	63
5	Terrence Sejnowski	87	25	Andrew Blake	62
6	Ian Foster	86	25	Lixia Zhang	62
6	Anil K. Jain	86	25	Ben Shneiderman	62
6	Jiawei Han	86	25	Jim Gray	62
6	Richard Lipton	86	25	Rajeev Motwani	62
7	Deborah Estrin	84	26	Christos H. Papadimitriou	61
7	David A. Patterson	84	26	Jitendra Malik	61
8	Michael I. Jordan	82	26	David J. DeWitt	61
9	Herbert A. Simon	81	26	Vipin Kumar	61
10	Thomas S. Huang	79	27	Mihir Bellare	60
11	John McCarthy	78	27	David Karger	60
11	Sebastian Thrun	78	27	Ion Stoica	60
12	Philip S. Yu	76	27	Oded Goldreich	60
12	David Culler	76	28	Ian Horrocks	59
13	David Haussler	75	28	John A. Stankovic	59
13	Michael Franklin	75	28	Didier Dubois	59
14	Tomaso Poggio	74	29	Henri Prade	58
15	Hari Balakrishnan	73	29	Sushil Jajodia	58
15	David E. Goldberg	73	29	Martin Abadi	58
15	John Mitchell	73	29	Moni Naor	58
16	Steven Salzberg	72	29	Carl Kesselman	58
16	Georgios Giannakis	72	28	Adrian Perrig	57
16	Jennifer Widom	72	28	Joseph M. Hellerstein	57
17	Don Towsley	71	28	Mihalis Yannakakis	57
17	Takeo Kanade	71	29	Bernhard Schölkopf	56
18	Andrew Zisserman	70	29	W. Bruce Croft	56
18	Kai Li	70	29	Leonidas J. Guibas	56
19	Richard Karp	68	29	Christian S. Jensen	56
19	Christos Faloutsos	68	29	Vern Paxson	56
20	Daphne Koller	67	29	Jon Kleinberg	56
21	Alex Pentland	66	30	Amit Sheth	55
21	Stanley Osher	66	30	Randy H. Katz	55
22	Robert Tarjan	65	30	Noga Alon	55

30	C. Lee Giles	55	36	Charles E. Perkins	50
30	Prabhakar Raghavan	55	37	William Dally	49
30	George Varghese	55	37	Raghu Ramakrishnan	49
30	Xin Yao	55	37	Eduardo Sontag	49
31	Steffen Staab	54	37	Rajeev Alur	49
31	Robert Kraut	54	37	Rajkumar Buyya	49
31	Shree Nayar	54	37	Giuseppe De Giacomo	49
31	David Dill	54	37	Avi Wigderson	49
31	Francisco Herrera	54	37	Pat Hanrahan	49
31	Jack Dongarra	54	37	Monica S. Lam	49
31	Olivier Faugeras	54	38	Jeffrey D. Ullman	48
31	Ramesh Govindan	54	38	Ronald Fagin	48
32	Sally Floyd	53	38	David Garlan	48
32	John Hennessy	53	38	Sara Kiesler	48
32	Thomas A. Henzinger	53	38	Hsinchun Chen	48
32	Henry Levy	53	38	Alberto Sangiovanni-Vincentelli	48
33	Demetri Terzopoulos	52	38	Gene Tsudik	48
33	Marco Dorigo	52	38	Jose Meseguer	48
33	Alon Halevy	52	39	Maurizio Lenzerini	47
33	Amir Pnueli	52	39	Dinesh Manocha	47
33	Tim Finin	52	39	Pavel Pevzner	47
33	H. V. Jagadish	52	39	Ken Kennedy	47
33	Craig Chambers	52	39	Serge Abiteboul	47
33	George Karypis	52	39	Hans-Peter Kriegel	47
34	Luca Cardelli	51	39	Victor Lesser	47
34	James Hendler	51	39	Heikki Mannila	47
34	Dieter Fensel	51	39	Moshe Y. Vardi	47
34	Saul Greenberg	51	39	Nancy Lynch	47
34	Pierre Baldi	51	40	Michael Stonebraker	46
34	Richard Baraniuk	51	40	M. Frans Kaashoek	46
34	David Harel	51	40	Donald E. Knuth	46
34	Geoffrey E. Hinton	51	40	Herbert Edelsbrunner	46
34	Tuomas Sandholm	51	40	Maja Mataric	46
34	Douglas C. Schmidt	51	41	Leslie Lamport	45
34	Zoubin Ghahramani	51	41	Rama Chellappa	45

34	Miron Livny	51	41	Mark Handley	45
35	Ronald Yager	50	41	Chunming Qiao	45
35	Adi Shamir	50	41	Eero P. Simoncelli	45
35	Amin Vahdat	50	41	Victor Basili	45

41	Nick McKeown	45	47	John Canny	39
41	Raymond Mooney	45	47	Salvatore Stolfo	39
41	Ian Witten	45	47	Mohammed Zaki	39
42	Henning Schulzrinne	44	47	Krithi Ramamritham	39
42	Baruch Awerbuch	44	47	Andrew Appel	39
42	Jim Kurose	44	48	Lotfi Zadeh	38
42	JJ Garcia Luna Aceves	44	48	Michael Pazzani	38
42	Stephanie Forrest	44	48	David Eppstein	38
42	Micha Sharir	44	48	John Stasko	38
42	Joachim Weickert	44	48	Willy Zwaenepoel	38
43	Jeffrey Vitter	43	48	Eitan Altman	38
43	Jason Cong	43	48	Stefano Ceri	38
43	Brad Myers	43	49	Kurt Keutzer	37
43	Gregory D. Abowd	43	49	Frank Pfenning	37
43	Kim G. Larsen	43	49	Teuvo Kohonen	37
43	Barry Boehm	43	49	Erkki Oja	37
43	Michael Kearns	43	49	Andrew S. Tanenbaum	37
43	Songwu Lu	43	50	Alan Yuille	36
43	Benjamin Pierce	43	50	Kathleen McKeown	36
44	Jon Crowcroft	42	50	Toby Walsh	36
44	Judea Pearl	42	50	Margaret Martonosi	36
44	Gaurav Sukhatme	42	51	Andrew B. Kahng	35
44	Rajeev Rastogi	42	51	Joel Saltz	35
44	Simon Peyton Jones	42	51	Leonard Kleinrock	35
44	Barbara Liskov	42	51	Gordon Plotkin	35
44	Paul Dourish	42	51	Aaron Bobick	35
44	Milind Tambe	42	51	Chandrajit Bajaj	35
45	Gaetano Borriello	41	51	Rina Dechter	35
45	David Salesin	41	51	Ahmed Elmagarmid	35
46	HongJiang Zhang	40	51	Joseph Goguen	35
46	Peter Brusilovsky	40	51	Miodrag Potkonjak	35

46	Lothar Thiele	40	51	Ingemar J. Cox	35
46	Rajiv Gupta	40	52	Alan Newell	34
46	Robin Milner	40	52	Richard Snodgrass	34
46	Ronald L. Rivest	40	52	Erol Gelenbe	34
46	Ricardo Baeza-Yates	40	52	Luis Gravano	34
46	Rafail Ostrovsky	40	52	Zhengyou Zhang	34
46	Philip Wadler	40	53	Pradeep K. Khosla	33
47	Frank van Harmelen	39	53	Steven Feiner	33

54	Murat Tekalp	32
54	Robbert van Renesse	32
54	Darrell Whitley	32
54	Michael Arbib	32
54	Alex Aiken	32
54	Deborah L. McGuinness	32
54	Edmund M. Clarke	32
55	Zbigniew Michalewicz	31
56	Carlo Zaniolo	30
57	Norman I. Badler	29
57	John Koza	29
58	Tom Leighton	28
59	Daniel Siewiorek	27
60	Nick Jennings	23
61	Alex Nicolau	19
62	Ben Bederson	14
63	Thomas W. Reps	10
64	Ron Fedkiw	6
65	Ed G. Coffman	1

Συμπεράσματα – Μελλοντική δουλειά

Η εκτίμηση της σημασίας του έργου ενός επιστήμονα είναι ένα πολύ σημαντικό θέμα για απονομές βραβείων, την πρόσληψη σε κάποιο πανεπιστημιακό ίδρυμα ή ερευνητικό κέντρο κτλ.. Αυτό το ζήτημα όπως είδαμε έχει λάβει μεγάλη προσοχή και όλο και περισσότεροι δείκτες προτείνονται για αυτό το σκοπό.

Μια βασική αδυναμία που συναντάμε στον H-index, το coterminal citation, επιλύεται και βελτιστοποιείται στον F-index, χωρίς όμως να μην κληρονομεί αδυναμίες του H-index.

Με την παρούσα εργασία καταφέραμε να δώσουμε συνέχεια σε αυτή την προσπάθεια, με την προσθήκη χρονικών περιορισμών και την εισαγωγή του δείκτη trend F-index. Μπορούμε να δούμε σε μια συγκεκριμένη περίοδο και περισσότερο στη σημερινή που μας ενδιαφέρει ποιων επιστημόνων τα άρθρα είναι που ελκύουν περισσότερο ενδιαφέρον και αξιοποιούνται περισσότερο η δουλειά τους από την επιστημονική κοινότητα.

Ως μελλοντική δουλειά χρειάζεται να ολοκληρωθεί η προσπάθεια αξιοποιώντας όλα τα είδη δεικτών που αναφέρουμε ως βελτιστοποιήσεις του F-index (contemporary, normalized, yearly, normalized yearly), στα πρότυπα του trend H-index.

Βιβλιογραφία

Σαχίνη Ε., Μάλλιου Ν., Χούσος Ν., Βιβλιομετρική Ανάλυση: Μεθοδολογική Προσέγγιση ΕΚΤ, Εθνικό Κέντρο Τεκμηρίωσης, 3^η έκδοση, Δεκέμβριος 2014

Garfield, E. (1964). Can citation indexing be automated? In M.E. Stevens, V.E. Giuliano, & L.B. Heilprin, (Eds.), *Proceedings of the Symposium on Statistical Association Methods for Mechanized Documentation* (pp. 189–192). Washington, DC: U.S. Government Printing Office.

HIRSCH, J. E., (2005) An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Sciences*, 102 (46) 16569–16572.

Katsaros, D., Akritidis, L., & Bozanis, P. (2009). The f index: Quantifying the impact of coterminal citations on scientists' ranking. *Journal of the American Society for Information Science & Technology*, 60(5), 1051–1056.

van Raan, A.F.J. (2008). Self-citation as an impact-reinforcing mechanism in the science system. *Journal of the American Society for Information Science and Technology*, 59, 1631–1643.

Schreiber, M. (2007). Self-citation corrections for the Hirsch index. *Euro-physics Letters*, 78(3), article number 30002.1-30002.6.

Sidiropoulos, A., Katsaros, D., & Manolopoulos, Y. (2007). Generalized Hirsch h-index for disclosing latent facts in citation networks. *Scientometrics*, 72(2), 253–280.