

ΠΑΝΕΠΙΣΤΗΜΙΟ ΚΡΗΤΗΣ

ΤΜΗΜΑ ΕΠΙΣΤΗΜΗΣ ΥΠΟΛΟΓΙΣΤΩΝ

ΠΑΡΟΥΣΙΑΣΗ / ΕΞΕΤΑΣΗ ΜΕΤΑΠΤΥΧΙΑΚΗΣ ΕΡΓΑΣΙΑΣ

**Κορνιλάκης Αλέξανδρος
Μεταπτυχιακός Φοιτητής**

**Τμήμα Επιστήμης Υπολογιστών, Πανεπιστήμιο Κρήτης
Επόπτης Μεταπτ. Εργασίας: Καθηγητής, Ε. Μαρκάτος**

Δευτέρα, 16/09/2019, 12:30

Αίθουσα Β106, Τμήμα Επιστήμης Υπολογιστών, Πανεπιστήμιο Κρήτης

“CHECK-IT: Ανίχνευση Ψευδών Ειδήσεων σε Πραγματικό Χρόνο”

ΠΕΡΙΛΗΨΗ

Τα τελευταία χρόνια, παρατηρούμε την άνοδο της παραπληροφόρησης στο Διαδίκτυο. Οι χρήστες διαρκώς πέφτουν θύματα ψευδών ειδήσεων και συμβάλλουν στη διάδοση τους, εν γνώσει τους ή μη. Η χρήση προπαγάνδας είναι ένα διόλου πρόσφατο φαινόμενο, αλλά ποτέ πριν δεν υπήρχε η κατάλληλη τεχνολογία για να διαδοθεί τόσο αποτελεσματικά. Η χρήση των κοινωνικών μέσων που έχει λάβει σημαντικές διαστάσεις στη ζωή μας την τελευταία δεκαετία, ώθησε πρακτικά την αύξηση της παραπληροφόρησης, συμπεριλαμβανομένης της σχετικής διάδοσης ψεύτικων ειδήσεων. Οι πυροβολισμοί ως αποτέλεσμα της θεωρίας συνωμοσίας «rizzagate» και το σκάνδαλο Cambridge Analytica, υποδεικνύουν πώς δεν πρέπει να λάβουμε αψήφιστα αυτή την άνοδο της παραπληροφόρησης.

Αρκετές πρόσφατες δουλειές προσπαθούν να περιορίσουν τις συνέπειες από την διάδοση των ψευδών ειδήσεων, εντοπίζοντας τις αυτόματα χρησιμοποιώντας τεχνικές τεχνητής νοημοσύνης, γνωστές λίστες, πληροφορίες από κοινωνικά δίκτυα κλπ.

Στην παρούσα μεταπτυχιακή εργασία παρουσιάζουμε το Check-It, ένα σύστημα που συνδυάζει μια ποικιλία σημάτων με στόχο την ταυτοποίηση των ψευδών ειδήσεων. Τέτοια σήματα περιλαμβάνουν τη φήμη του ατόμου (λογαριασμού) που δημοσιεύει τις ειδήσεις, τη φήμη της ιστοσελίδας όπου φιλοξενούνται τα νέα, τα γλωσσικά στοιχεία που χαρακτηρίζουν ένα ψεύτικο ειδησεογραφικό άρθρο καθώς και το ίδιο το περιεχόμενο του άρθρου.

Χρησιμοποιώντας μια προσέγγιση βαθιάς μάθησης συνδυάζουμε όλα αυτά τα χαρακτηριστικά για τον υπολογισμό μιας έγκαιρης και ακριβούς ένδειξης. Το Check-It έχει αναπτυχθεί ως ένα plugin για προγράμματα περιήγησης ιστού με στόχο την αποτελεσματική και έγκαιρη ανίχνευση ψεύτικων ειδήσεων, ενώ παράλληλα σέβεται το ιδιωτικό απόρρητο των χρηστών. Οι απαιτήσεις που επιβάλαμε για τον σχεδιασμό του Check-It είναι η συμμόρφωση με το GDPR, η έλλειψη ανακριβών απαντήσεων, ο μικρός χρόνος απόκρισης και ο μικρός υπολογιστικός φόρτος. Για να υλοποιήσουμε το plugin μας έχουμε χρησιμοποιήσει βιβλιοθήκες γραμμένες σε JavaScript, όπως οι Minhash.js και TensorFlow.js.

Στην παρούσα εργασία παρουσιάζουμε το σχεδιασμό, την υλοποίηση και την αξιολόγηση της απόδοσης του Check-It. Τα πειραματικά αποτελέσματα πάνω σε γνωστά σύνολα δεδομένων δείχνουν ότι το Check-It συνιστά βελτίωση των σύγχρονων σχετικών μεθόδων, ενώ ταυτόχρονα επιτυγχάνει ακρίβεια της τάξης του 93%. Επιπλέον, ορίζουμε ορισμένες κατευθύνσεις που μπορούν να ακολουθήσουν οι μελλοντικές εκδόσεις του Check-It.

Alexandros Kornilakis

M.Sc. Thesis

Computer Science Department

University of Crete

Master's Thesis Supervisor: Professor, E. Markatos

Monday, 16/09/2019, 12:30

Room B106, Computer Science Dept., University of Crete

“CHECK-IT: Real Time Detection of Fake News”

ABSTRACT

Over the past few years, we have been witnessing the rise of misinformation on the Internet. People fall victims of fake news continuously and contribute to their propagation knowingly or inadvertently. The use of propaganda is indeed ancient, but never before has there been the technology to so effectively disseminate it. The social media engagement that has swept our lives over the past decade practically exploded the proliferation of misinformation, including the associated distribution of fake news. The ‘pizzagate shooting’ incident and the Cambridge Analytica scandal indicate that we should not take this rise of misinformation lightly.

Many recent efforts seek to reduce the damage caused by fake news by identifying them automatically with artificial intelligence techniques, using signals from domain flag-lists, online social networks, etc.

In this thesis, we present Check-It, a system that combines a variety of signals into a pipeline for fake news identification. Such signals include the reputation of the person (account) posting the news, the reputation of the website where the news is hosted, the linguistic features that characterize a fake news article as well as the article’s content per se.

Using a deep learning approach, we combine all these features towards providing a rating that is timely and accurate. Check-It is developed as a web browser plugin with the objective of efficient and timely fake news detection while respecting user privacy. The requirements we considered when designing Check-It is GDPR compliant, highly confident identification, low response time and lightweight computation. To implement our plugin, we have used pure JavaScript frameworks, like Minhash.js and TensorFlow.js.

In this thesis, we present the design, implementation, and performance evaluation of Check-It. Experimental results show that it outperforms state-of-the-art methods on commonly used datasets while achieving an accuracy of 93%. Furthermore, we provide some directions that can guide future versions of Check-It.