Ανοικτό Πανεπιστήμιο Κύπρου

Σχολή Θετικών και Εφαρμοσμένων Επιστημών

Μεταπτυχιακό Πρόγραμμα Σπουδών
Πληροφοριακά και Επικοινωνιακά Συστήματα

Μεταπτυχιακή Διατριβή

Χρήση Έξυπνων Εργαλείων Ανοιχτού Κώδικα για την Παρακολούθηση Στόχων μέσω των Social Networks Twitter και Instagram.

Κωνσταντίνος Τζοβελέκης

Επιβλέπων Καθηγητής
Δρ. Σταύρος Σιαηλής

Δεκέμβριος 2015

Ανοικτό Πανεπιστήμιο Κύπρου

Σχολή Θετικών και Εφαρμοσμένων Επιστημών

Μεταπτυχιακό Πρόγραμμα Σπουδών

Πληροφοριακά και Επικοινωνιακά Συστήματα

Μεταπτυχιακή Διατριβή

Χρήση Έξυπνων Εργαλείων Ανοιχτού Κώδικα για την Παρακολούθηση Στόχων μέσω των Social Networks Twitter και Instagram

Κωνσταντίνος Τζοβελέκης

Επιβλέπων Καθηγητής
Δρ. Σταύρος Σιαηλής

Η παρούσα μεταπτυχιακή διατριβή υποβλήθηκε προς μερική εκπλήρωση των απαιτήσεων για απόκτηση μεταπτυχιακού τίτλου σπουδών στα Πληροφοριακά Συστήματα από τη Σχολή Θετικών και Εφαρμοσμένων Επιστημών του Ανοικτού Πανεπιστημίου Κύπρου.

Δεκέμβριος 2015

ΛΕΥΚΗ ΣΕΛΙΔΑ

Περίληψη

Τα οφέλη της τεχνολογίας, μας περιτριγυρίζουν. Τα αυτοκίνητα και τα κινητά τηλέφωνα μας, περιλαμβάνουν συστήματα εντοπισμού θέσης, έτσι ώστε να μην χανόμαστε. Μπορούμε να πληρώνουμε τους λογαριασμούς μας με το πάτημα ενός κουμπιού στο Internet. Τα κινητά μας τηλέφωνα έχουν εξελιγχθεί σε μικροσκοπικούς υπολογιστές, βοηθώντας μας σε όλες τις καθημερινές μας ανάγκες. Βιώνουμε την εποχή της τεχνολογίας και της μαζικής ανταλλαγής πληροφοριών, με τις πλατφόρμες κοινωνικής δικτύωσης να παίζουν βασικό ρόλο στην αλλαγή του τρόπου της επικοινωνίας μας. Ο καθημερινός όγκος των πληροφοριών που διαμοιράζεται είναι τεράστιος, ως αποτέλεσμα της ανάγκης μας για επικοινωνία και αυτό προβολή. Αναπόφευκτα η χρήση της τεχνολογίας, οδηγεί στην αποκάλυψη προσωπικών δεδομένων από την πλευρά των χρηστών. Κάνοντας διάκριση ανάμεσα στο ίδιο το μέσο (πλατφόρμες κοινωνικής δικτύωσης) και στον τρόπο χρήσης τους από τους ίδιους τους χρήστες, θέτουμε το ερευνητικό ερώτημα του ενδεχομένου άντλησης προσωπικών στοιχείων με αυτοματοποιημένο, μη ανιχνεύσιμο τρόπο από έναν κακόβουλο χρήστη, έναντι ενός ανυποψίαστου προσώπου, χρήστη των πλατφορμών κοινωνικής δικτύωσης.

Σαν απάντηση στο παραπάνω ερευνητικό ερώτημα, η παρούσα μεταπτυχιακή διατριβή πραγματεύεται την άντληση στοιχείων με αυτοματοποιημένο τρόπο, από δύο δημοφιλή κοινωνικά δίκτυα Twitter και Instagram, χωρίς την συγκατάθεση ή την ενημέρωσή του τελικού χρήστη – στόχου. Τα αποτελέσματα δείχνουν ότι η δραστηριότητα μας στα κοινωνικά δίκτυα μπορεί να μπορεί να αποτελέσει αντικείμενο εκμετάλλευσης από κάποιον κακόβουλο χρήστη ο οποίος μέσω τεχνικών OSINT συγκεντρώνει ευαίσθητα προσωπικά δεδομένα (θέση χρήστη μέσω tweets / instas από το smartphone μας). Αξίζει να αναφερθεί ότι η γεωγραφική θέση της δραστηρίοτητας του χρήστη στόχου, που έλαβε χώρα στα δύο συγκεκριμένα κοινωνικά δίκτυα, αποκαλύπτεται από τα δεδομένα GPS του κινητού τηλεφώνου, ενώ στην συνέχεια γίνεται απεικόνιση των σημείων δραστηρίοτητας σε χάρτη τύπου Google maps, συνδέοντας μάλιστα τα σημεία χρονολογικά. Με αυτόν τον τρόπο ο κακόβουλος χρήστης μπορεί να παρακολουθήσει την θέση του ανυποψίαστου προσώπου και ακόμα να προβλέψει την μελλοντική του θέση.

Summary

The benefits of technology are all around us. Our cars and cell phones have navigation systems to ensure that we do not get lost. We can pay our bills at the touch of a button on the internet. Our cell phones evolved into tiny little computers helping us in all of our daily tasks. We are living in the era of technology and massive information exchange and social networks are very important part as they drastically changed how we communicate. The daily amount of information shared is tremendous as a result of the need for communication and self-promotion of ourselves. The use of this technology, inevitably leads to disclosure of personal data on the part of users. Making a distinction between the medium itself (social networking platform) and how it is used by users, we set the research question of how possible is a malicious user to retrieve personal data in an automated undetectable way and use them against an unsuspicious person that uses social media.

In regard to previous question, this thesis had intended to detect the possibility of retrieving information in an automated way of two popular social networks, Twitter and Instagram, without the consent or informing the end user - target. The results shown that our activity on social networks, can be exploited by a malicious user using OSINT techniques and can collect sensitive personal data (user location via tweets / instas from smartphone). It is worth to mention that the location activity took place in the two particular social networks, through the targeted user smartphone, reveals the GPS coordinates of the user activity and this was reproduced in a Google Maps type map where we also placed connecting points chronologically. This way a malicious user can track an unsuspicious person and it is possible to predict future location.

Ευχαριστίες

Θα ήθελα να εκφράσω τις θερμές μου ευχαριστίες στον επιβλέποντα καθηγητή μου Δρ. Σταύρο Σιαηλή για τη συνεχή καθοδήγηση, τις πολύτιμες πληροφορίες και ενδείξεις που μου έδωσε καθ' όλη τη διάρκεια για τη εκπλήρωση της διατριβής.

Επιπρόσθετα θα ήθελα να ευχαριστήσω την οικογένειά μου για την συνεχή υποστήριξη τα τελευταία δύο χρόνια, ιδιαίτερα την σύζυγό μου Κάλλια και την εννιάχρονη κόρη μου Μυρσίνη.

Περιεχόμενα

Κεφάλαιο 1
Εισαγωγή

Η πραγματική πρόκληση σήμερα δεν είναι πλέον η πρόσβαση σε ενημέρωση αλλά η επεξεργασία και η ανάλυση των εν λόγω πληροφοριών. Στην εποχή μας συναντάμε υπερπληθώρα πληροφοριών, ενώ σοβαρότερο πρόβλημα αποτελεί η αξιολόγηση και η αξιοπιστία των πληροφοριών.

Η εμφάνιση του Web 2.0 έχει συμβάλει στη μετατροπή ενός μέσου χρήστη του διαδικτύου από παθητικό αναγνώστη σε έναν διαδραστικό χρήστη. Οι χρήστες έχουν πλέον τη δυνατότητα να αλληλοεπιδρούν με άλλους ανθρώπους, να δημιουργούν, ή να ανταλλάσσουν πληροφορίες και απόψεις, καθώς επίσης να εκφράσουν τις απόψεις τους σε εικονικές κοινότητες. Όλα τα παραπάνω που προσφέρονται μέσα από το Web 2.0, αναφέρονται ως Social Media [Kaplan, 2010:59]. Έτσι μέσα από τα Social Media μπορούν οι χρήστες να δημιουργήσουν την προσωπική τους ταυτότητα στον ψηφιακό κόσμο, παρουσιάζοντας ταυτόχρονα πτυχές της προσωπικής τους ζωής.

Οι χρήστες πλέον τείνουν να συμμετέχουν στα Social Media για διάφορους λόγους, όπως επαγγελματικούς, διασκέδασης, αλλά και επικοινωνίας. Η παρατήρηση αυτής της συνεχούς χρήσης των κοινωνικών δικτύων έχει οδηγήσει σε έρευνες που έχουν δείξει ότι τα άτομα τείνουν να μεταφέρουν την προσωπικότητα τους στο διαδίκτυο [Amichai-Hamburger, 2010:1290].

Χρησιμοποιώντας τεχνικές OSINT [Hulnick, 2010:58], μπορεί κανείς εύκολα να παρατηρήσει τα διάφορα πρότυπα αλλά και τον τρόπο χρήσης του κάθε χρήστη των μέσων κοινωνικής δικτύωσης. Επιπλέον, αυτά τα δεδομένα που εξάγονται μέσω των τεχνικών OSINT μπορούν να χρησιμοποιηθούν για σκοπούς, όπως αυτό της στοχευμένης διαφήμισης, της καταγραφής του προφίλ της προσωπικότητας του χρήστη, της

πρόβλεψης της συμπεριφοράς του, αλλά και της θέσης όπου γίνεται χρήση του εκάστοτε μέσου κοινωνικής δικτύωσης.

Ένα σημαντικό χαρακτηριστικό των μέσων κοινωνικής δικτύωσης είναι ότι τα δεδομένα προσωπικού χαρακτήρα, είναι διαθέσιμα για την ανίχνευση και την επεξεργασία, ακόμη και χωρίς τη συγκατάθεση του χρήστη. Παρά το γεγονός ότι, η επεξεργασία των δεδομένων αυτών μπορεί να χρησιμοποιηθεί για αθώους "σκοπούς", όπως στοχευμένο μάρκετινγκ, ή έλεγχος της εμπειρίας του χρήστη, εντούτοις μπορεί να παραβιάσει την ιδιωτική ζωή του χρήστη. Η δυνατότητα διενέργειας αυτοματοποιημένων ψυχομετρικών αξιολογήσεων καθώς επίσης η αποκάλυψη προσωπικών δεδομένων σχετικά με τις πολιτικές και θρησκευτικές πεποιθήσεις, θέτει σε κίνδυνο την προστασία της ιδιωτικής ζωής των χρηστών και είναι σημαντικό να θεσμοθετηθεί νομοθεσία, με σκοπό την προστασία του χρήστη στο διαδίκτυο [Bean, 2007:242].

Τα εργαλεία που χρησιμοποιούνται στις τεχνικές OSINT αλλά και την επεξεργασία των διαδικτυακών δεδομένων του χρήστη μπορούν να διακριθούν σε αυτά όπου υπάρχει προοπτική παραβίασης των δεδομένων αναφορικά με την ιδιωτικότητα του χρήστη και σε αυτά που υπάρχει βελτιωμένη και απρόσωπη εμπειρία του χρήστη. Στην πρώτη περίπτωση, τα δεδομένα του χρήστη επεξεργάζονται, χωρίς τη συγκατάθεσή του, ενώ στην δεύτερη ο χρήστης είναι εφοδιασμένος με ένα πιο φιλικό περιβάλλον προσαρμοσμένο στις προτιμήσεις του.

Η εργασία αυτή έχει ως σκοπό να ερευνήσει τις προκλήσεις του Open Source Intelligence (OSINT) στα Social Media. Στην συνέχεια θα γίνει αναφορά των τεχνικών που χρησιμοποιούνται για το φιλτράρισμα των δεδομένων που διακινούνται μέσα από τις πλατφόρμες κοινωνικής δικτύωσης και συγκεκριμένα των Twitter και Instagram, ενώ στην συνέχεια παρουσιάζονται τρόποι εξαγωγής δεδομένων τοποθεσίας των χρηστών τους. Η εργασία χωρίζεται στις παρακάτω θεματικές ενότητες (κεφάλαια), την πρώτη ενότητα "Εισαγωγή" όπου γίνεται μια αναφορά σχετικά με τα μέσα κοινωνικής δικτύωσης, την δεύτερη ενότητα "Θεωρία OSINT" μέσα από την οποία γίνεται ανάλυση των ποιο διαδεδομένων τεχνικών "OSINT" καθώς επίσης παρουσιάζονται τα αρνητικά και θετικά της κάθε τεχνικής. Στο τρίτο κεφάλαιο γίνεται μία επισκόπηση των τεχνικών που χρησιμοποιούν οι hackers για να παραμείνουν ανώνυμοι στο διαδίκτυο σήμερα.

Στα επόμενα τρία κεφάλαια 4,5 και 6 γίνεται ανάλυση των Διεπαφών Προγραμματισμού Εφαρμογών (API) των Twitter, Instagram και Google, παραθέτοντας τις βασικές λειτουργίες των παραπάνω εφαρμογών. Στο έβδομο κεφάλαιο αναλύεται το "script" δηλαδή το πρόγραμμα και πως αυτό εξάγει τα δεδομένα τοποθεσίας από το Twitter και το Instagram. Στη όγδοη ενότητα αναλύονται τα νομικά και ηθικά ζητήματα που προκύπτουν από την εξαγωγή δεδομένων από τα κοινωνικά μέσα δικτύωσης, ενώ στην ένατη ενότητα παρουσιάζονται οι τρόποι προστασίας της ιδιωτικότητας των χρηστών των μέσων κοινωνικής δικτύωσης.

Στο δέκατο κεφάλαιο έχουμε τα συμπεράσματα που απορρέουν από την εργασία αλλά και προτάσεις οι οποίες μπορούν να χρησιμοποιηθούν σε επόμενες έρευνες. Στο Παράρτημα Α γίνεται παρουσίαση του κώδικά που χρησιμοποιήθηκε στην έρευνα, ενώ στο Παράρτημα Β παραθέτονται παραδείγματα της χρήσης του κώδικά εξαγωγής δεδομένων τοποθεσίας.

Η καινοτομία της συγκεκριμένης διατριβής εντοπίζεται στο γεγονός ότι επικεντρώνεται σε δύο μεγάλα κοινωνικά δίκτυα με mobile first λογική και συνδυάζει τα γεωγραφικά σημεία δραστηριότητας του χρήστη – στόχου απο αυτά, σε έναν κοινό χάρτη με διανύσματα κατά χρονολογική σειρά. Επιπλέον προσφέρει μια μεθοδολογία, με βάση την οποία παραμένουμε μη ανιχνεύσιμοι και πραγματοποιούμε την επίθεση στον χρήστη – στόχο ανώνυμα.

Κεφάλαιο 2
OSINT

2.1 Τι είναι OSINT

Ο όρος Open Source Intelligence (OSINT) κατά τον Arthuer S. Hulnick αναφέρεται ως μια πηγή πληροφοριών καταλόγου [Hulnick, 2010:58]. Ακόμη το OSINT ορίζεται ως μη διαβαθμισμένες πληροφορίες που λαμβάνονται από οποιαδήποτε διαθέσιμη πηγή σε έντυπη, ηλεκτρονική, ακόμα και λεκτική μορφή. Ραδιόφωνο, τηλεόραση, εφημερίδες, περιοδικά, διαδίκτυο, εμπορικές βάσεις δεδομένων, βίντεο, όλα εμπίπτουν στην κατηγορία αυτή και μπορούν να αξιοποιηθούν με επιτυχία για τη συγκέντρωση πληροφοριών [Hulnick, 2010:63, Bean, 2007:243].

Η πραγματική πρόκληση για το OSINT δεν είναι η αναζήτηση και η ανάκτηση των σχετικών πληροφοριών, αλλά το φιλτράρισμα, η κατηγοριοποίηση και η επεξεργασία των τεράστιων ποσοτήτων δεδομένων προκειμένου ο αναλυτής να εξάγει την χρήσιμη "πληροφορία". Δηλαδή για να λειτουργήσει σωστά η διαδικασία απαιτείται η αποθήκευση των μεταδεδομένων που προέρχονται από κάποια τεχνική OSINT αλλά και από πρόγραμμα ανάλυσης αυτών. Η αποτελεσματική δημιουργία ευρετηρίου και αρχειοθέτησης των πληροφοριών, απαιτεί βασική γνώση των βάσεων πληροφοριών που καλύπτει η περιοχή μελέτης. Αυτά τα είδη βάσεων πληροφοριών πολλές φορές αναφέρονται αόριστα ως μια οντολογία ή ως μια ταξινόμηση. Μια οντολογία περιγράφει τις έννοιες, τους όρους και τις οντότητες καθώς επίσης τους σχετικούς κανόνες, για το πώς αλληλοεπιδρούν και σχετίζονται μεταξύ τους. Μια ταξινόμηση, είναι απλώς μια ιεράρχηση των θεμάτων, υποκατηγοριών και μεμονωμένων υπάρξεων που περιγράφουν έναν τομέα. Έτσι, για παράδειγμα, μια ταξινόμηση που περιγράφει το Ευρωπαϊκό Κοινοβούλιο θα αποτελείται από το εκάστοτε "κόμμα", το οποίο με τη σειρά τους αποτελείται από μεμονωμένες ομάδες που αποτελούνται από μεμονωμένα άτομα που βρίσκονται σε διαφορετικά κράτη της Ευρώπης. Οι βάσεις πληροφόρησης που προέρχονται από τεχνικές OSINT παρέχουν το πλαίσιο που επιτρέπει τη διασταύρωση

των εισερχόμενων εκθέσεων με δομημένο πλαίσιο μέσα στο χώρο και το χρόνο. Οι λεπτομέρειες της κάθε βάσης γνώσεων εξαρτώνται από το συγκεκριμένο υπό μελέτη τομέα.

Αυτό το είδος των δεδομένων είναι εξαιρετικά χρήσιμα και συχνά είναι το μόνο μέσο διείσδυσης σε μυστικά δίκτυα. Η διαδικασία αρχίζει με στοιχεία ανοικτού κώδικα Open Source Data (OSD), τις πρώτες πληροφορίες από την πρωτογενή πηγή και στη συνέχεια πρέπει να συναρμολογηθούν μέσα από μια διαδικασία επεξεργασίας έτσι ώστε να φιλτραριστούν και να επικυρωθούν [Bean, 2007:244, Lahneman, 2010:203].

Αυτό στη συνέχεια οδηγεί σε πληροφορίες ανοιχτού κώδικα Open Source Information (OSIF), που μπορούν να διαδοθούν ως άρθρα σε εφημερίδες, βιβλία, τηλεοπτικές και ραδιοφωνικές εκπομπές και στο διαδίκτυο. Έτσι μια πηγή πληροφοριών καταλόγου (OSINT) δημιουργείται μόνο εφόσον η πληροφορία ανοιχτού κώδικα έχει ανακαλυφθεί, αναλυθεί και διαδοθεί σε ένα επιλεγμένο ακροατήριο σε σχέση με ένα συγκεκριμένο ζήτημα [Lahneman, 2010:204].

2.2 Social Media

Τα μέσα κοινωνικής δικτύωσης είναι ένας όρος που χρησιμοποιείται ευρέως για να περιγράψει οποιοδήποτε αριθμό των τεχνολογικών συστημάτων που σχετίζονται με τη συνεργασία και την κοινότητα [Joosten, 2012:183]. Ενώ φαίνεται ότι ένας συγκεκριμένος ορισμός μπορεί να είναι ασαφής [Best, 2011:59], τα social media συχνά περιγράφονται σύμφωνα με την χρήση τους. Οι ιστότοποι κοινωνικής δικτύωσης, ιστολόγια, wikis, πλατφόρμες πολυμέσων, παιχνίδια εικονικού κόσμου, είναι μεταξύ των εφαρμογών συνήθως που περιλαμβάνονται στα social media [Ignat, 2005:259]. Για να περιορίσουμε το εύρος, οι ιστότοποι κοινωνικής δικτύωσης που επιλέχθηκαν ως οι επικρατέστεροι των μέσων κοινωνικής δικτύωσης, είναι το Facebook, MySpace, LinkedIn, Twitter. Εναλλακτική ορολογία όπως είναι η κοινωνική δικτύωση ή online κοινωνικά δίκτυα, είναι web-based υπηρεσίες που επιτρέπουν στους χρήστες να κάνουν προσωπικά προφίλ, να δημιουργούν περιεχόμενο, και να μοιράζονται τα μηνύματα με άλλους χρήστες του συστήματος [Ellison, 2007:212]. Το έργο του προσδιορισμού των κοινωνικών μέσων μαζικής ενημέρωσης γίνεται πιο δύσκολο από το γεγονός ότι είναι συνεχώς σε κατάσταση αλλαγής, καθώς η κάθε πλατφόρμα κοινωνικής δικτύωσης προσπαθεί να

δημιουργήσει νέα ή βελτιωμένα χαρακτηριστικά που θα καλύψουν τις ανάγκες των χρηστών. Μια σύντομη περιγραφή των λειτουργιών του Facebook απεικονίζει τις πιο πρόσφατες δυνατότητες δικτύων κοινωνικής δικτύωσης. Στο εσωτερικό του Facebook, οι χρήστες μπορούν να στέλνουν μηνύματα, να προσθέτουν νέους φίλους, να ενημερώνουν το προσωπικό προφίλ, να συμμετέχουν σε ομάδες, να αναπτύσσουν εφαρμογές, καθώς επίσης να ενημερώνονται για τους άλλους χρήστες μέσω των online προφίλ τους [Quan-Haase, 2010:254].

2.2.1 Facebook

Το Facebook μπορεί να θεωρηθεί ως το μεγαλύτερο μέσο κοινωνικό δίκτυο παγκοσμίως. Αναπτύχθηκε το 2004 από τον Mark Zuckerberg, και σύμφωνα με τον Lenhart et al είναι η ''κυρίαρχη'' πλατφόρμα κοινωνικής δικτύωσης [Lenhart, 2010:23]. Πρόσφατες μελέτες όπως αυτή που έγινε στο Harvard το 2011, έδειξε ότι σε ποσοστό πάνω από το 90% των φοιτητών, είχε λογαριασμό στο Facebook [Coughlan, 2015]. Τον Οκτώβριο του 2014 έφτασε να έχει 1,44 δισεκατομμύρια μηνιαίους ενεργούς χρήστες [Facebook, 2015].

2.2.2 LinkedIn

Το LinkedIn χρησιμοποιείται κυρίως για επαγγελματική δικτύωση, καθώς οι χρήστες τους συνήθως συσχετίζονται με τους υπόλοιπους στην εργασία τους, διατηρώντας μια λίστα επαφών για τους ανθρώπους που γνωρίζουν και εμπιστεύονται. Ο συντελεστής εμπιστοσύνης είναι μια σημαντική έννοια σε αυτήν την πλατφόρμα κοινωνικής δικτύωσης, καθώς η σύνδεση με τους άλλους απαιτεί είτε μια προϋπάρχουσα σχέση ή κάποια αμοιβαία επαφή [Papacharissi, 2009:203].

2.2.3 Twitter

Το Twitter είναι μία online πλατφόρμα κοινωνικής δικτύωσης αλλά και μια micro-blogging υπηρεσία, η οποία επιτρέπει στους εγγεγραμμένους χρήστες να δημοσιεύουν αλλά και να διαβάζουν σύντομα μηνύματα (γνωστά ως tweets) άλλων χρηστών της υπηρεσίας. Το Twitter μπορεί να χαρακτηριστεί σαν ένα δίκτυο ενημέρωσης πραγματικού χρόνου [Du, 2006:792].

Το Twitter υλοποιήθηκε με βάση την γλώσσα προγραμματισμού Ruby και διαθέτει το δικό του API (Application Programming Interface). Η υπηρεσία δημιουργήθηκε το 2005 από τον Jack Dorsey καθώς σκέφτηκε ότι θα ήταν πολύ ενδιαφέρον εάν μπορούσε να γνωρίζει τι κάνουν οι φίλοι του. Η πρώτη έκδοση την υπηρεσίας υλοποιήθηκε σε διάστημα μόλις δυο εβδομάδων ενώ η πρώτη επίσημη εμφάνιση του στο παγκόσμιο ιστό έγινε τον Αύγουστο του 2006. Η υπηρεσία έγινε πολύ σύντομα δημοφιλής με αποτέλεσμα τον Μάιο 2007 να ιδρυθεί η εταιρία "Twitter Incorporated". Σύμφωνα με το Twitter eMarketer [Twitter, 2015], τον Ιανουάριο του 2014 η υπηρεσία είχε εγγεγραμμένους περισσότερους από 645,750,000 χρήστες, από τους οποίους 135.000 συνδέονται σε καθημερινή βάση ενώ καθημερινά ανταλλάσσονται περισσότερα από 60 εκατομμύρια tweets. Τα tweets μπορούν να χαρακτηριστούν ως ηλεκτρονικά μηνύματα παρόμοια των Short Message Service (SMS) με την διαφορά ότι τα tweets έχουν δημόσια κοινοποίηση στην πλατφόρμα τους. Η λειτουργία αυτής της υπηρεσίας βασίζεται στο ότι ο χρήστης έχει τη δυνατότητα να ενημερώσει σχετικά με την κατάσταση του, άλλους χρήστες, δηλαδή τις σκέψεις , επιθυμίες, προβληματισμούς αλλά και το που βρίσκεται αυτή την στιγμή. Ο κάθε χρήστης έχει την δυνατότητα να παρακολουθεί τα μηνύματα άλλων χρηστών καθώς επίσης να τα σχολιάζει. Για να μπορέσουν οι χρήστες του Twitter να λειτουργήσουν ως ένα δίκτυο ανθρώπων, θα πρέπει να δημιουργήσουν το κύκλο τους. Στο Twitter υπάρχουν οι followers, αυτοί δηλαδή που ακολουθούν ένα χρήστη και ειδοποιούνται για κάθε μήνυμα και οι following αυτοί που ακολουθεί ο χρήστης και ενημερώνονται για τις αναρτήσεις τους.

2.2.4 Geolocation

Η ικανότητα γεωγραφικής κατάταξης μεγάλου όγκου δεδομένων Twitter είναι πολύτιμη για αρκετές ερευνητικές κατευθύνσεις. Πράγματι, η ανάλυση γεωγραφικών δεδομένων σε social media έχει αποδειχθεί χρήσιμη για την κατανόηση των περιφερειακών τάσεων όπως της γρίπης [Paul, 2011:266], γλωσσικών πρότυπων [Mocanu, 2013:7], εκλογικής πρόβλεψης [Tumasjan, 2010:180], των κοινωνικών αναταραχών [Compton, 2014:395], και την αντιμετώπιση των καταστροφών [Mandel, 2012:29]. Αυτές οι προσεγγίσεις, ωστόσο, εξαρτώνται από τις φυσικές θέσεις των χρηστών του Twitter, που είναι συνήθως δεδομένα, ελάχιστα διαθέσιμα δημόσια. Είναι ενδιαφέρον ότι, σε πρόσφατες έρευνες έχει διαπιστωθεί ότι η δημιουργία απευθείας κοινωνικών δεσμών σχηματίζονται

πιο συχνά σε μικρές γεωγραφικές αποστάσεις [Takhteyev, 2012:76, Mok, 2010:2750]. Εξαιτίας αυτού, είναι δυνατόν να προσεγγίσουμε τη θέση ενός χρήστη Twitter εξετάζοντας δημόσια γνωστές θέσεις των online φίλους τους [Jurgens, 2013:275, [Yamaguchi, 2013:228]. Η χρήση ανάλυσης των κοινωνικών δικτύων για την επίλυση προβλημάτων εύρεσης της τοποθεσίας βασίζεται μόνο στα δημόσια μεταδεδομένα του Twitter, η οποία, παρέχει πολλά πλεονεκτήματα σε σχέση με τις προσεγγίσεις που βασίζονται στο περιεχόμενο. Τα μεγαλύτερα και πιο ακριβή αποτελέσματα εύρεσης γεωγραφικών τοποθεσιών (geolocation) έκαναν χρήση της ανάλυσης των κοινωνικών δικτύων [Jurgens, 2013:278, Backstrom, 2010:65].

Geolocation [Memon, 2014:280, Chandra, 2011:840] είναι μια τεχνική η οποία προσπαθεί να εντοπίσει την θέση στην οποία δημιουργήθηκε ή δημοσιεύθηκε περιεχόμενο μια πλατφόρμας κοινωνικής δικτύωσης. Τα ονόματα των τοποθεσιών προσδιορίζονται μέσα από ένα ασαφή (fuzzy) τρόπο αναζήτησης δηλαδή μέσα από ένα 'λεξικό' από τοποθεσίες παρέχοντας στοιχεία σχετικά με το μήκος και το πλάτος της. Αυτό είναι πιο περίπλοκο από ό, τι φαίνεται, διότι διάφορες τοποθεσίες μπορούν να αναφερθούν αλλά να μην έχουν καμία σχέση με την τοποθεσία του δημοσιευμένου κείμενου [Compton, 2014:395 ,Valkanas, 2012:830]. Συνήθως ένα στατιστικό μέτρο που χρησιμοποιείται είναι να επιλέγει η πιο κοινή ονομασία θέσης σύμφωνα με τις εμφανίσεις της περιοχής και της χώρας που την περιέχουν. Η ακρίβεια των τεχνικών geolocation αυξάνεται όταν εφαρμόζεται σε ένα σύμπλεγμα από άρθρα που αναφέρονται στο ίδιο γεγονός. Καλύτερη ακρίβεια για την εύρεσης της τοποθεσίας επιτυγχάνεται με τον προσδιορισμό των βασικών φράσεων που περιέχει το όνομα του τόπου, ή απλά αναλύοντας τον τίτλο του άρθρου [Valkanas, 2012:831, Dhavase, 2014:23, Rakesh, 2013:1443].

Κεφάλαιο 3
Διατήρηση Ανωνυμίας

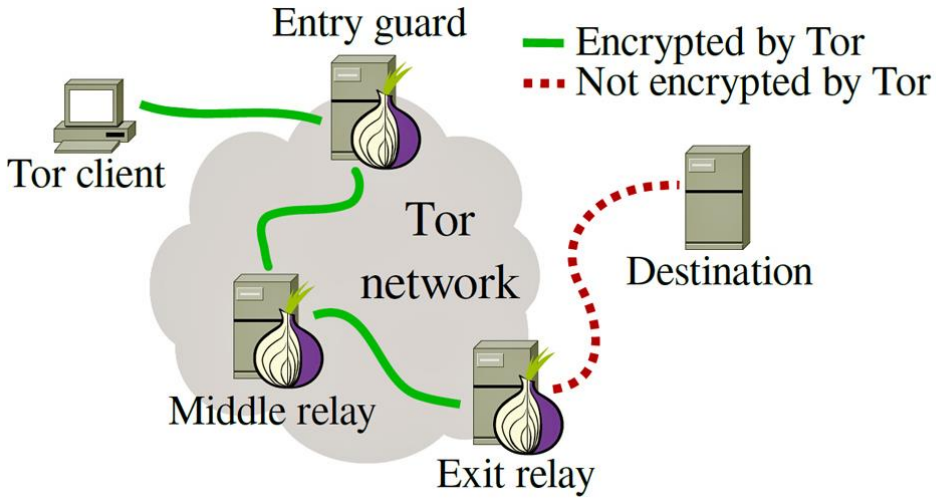
Στην προσπάθειά μας να συγκεντρώσουμε δεδομένα από τα social media όσο το δυνατόν πιο ανώνυμα και χωρίς να αφήσουμε ίχνη, στραφήκαμε σε «γκρίζες» μεθόδους και εργαλεία τα οποία δεν θα πρόδιδαν την ταυτότητα και τους σκοπούς μας. Προσπαθήσαμε δηλαδή να σκεφτούμε και να πράξουμε ως κακόβουλος χρήστης (hacker).

Οι προσπάθειές μας θα επικεντρωθούν σε δύο άξονες: αφενός την ανωνυμία και αφετέρου τη μη ανισχνευσιμότητα από τον στόχο χρήστη μας ή από την social πλατφόρμα στην οποία βασιζόμαστε. Όλες οι ενέργειές μας δεν θα πρέπει να φανερώνουν την ταυτότητα μας ή να αφήνουν σημαντικά ίχνη. Η συγκεκριμένη μεθοδολογία είχε επιτυχή αποτελέσματα την περίοδο που εφαρμόστηκε για τον σκοπό της συγγραφής της συγκεκριμένης πτυχιακής εργασίας (Αύγουστος – Οκτώβριος 2015) και έχοντας ως αφετηρία την Ελλάδα.

Θα πρέπει να αναφέρουμε ότι σχετικά με την μεθοδολογία διατήρησης ανωνυμίας που καταγράφεται στην συγκεκριμένη διατριβή, έγινε προσπάθεια εφαρμογής της με βάση γνωστές νόμιμες ή γκρίζες μεθόδους απόκρυψης της ταυτότητας μας. Δεν θα μπορούσαμε να ισχυριστούμε ότι είναι 100% μη ανιχνεύσιμη, καθότι πάντα οι δικτυακές συνδέσεις αφήνουν πίσω τους στοιχεία που τελικά θα μπορούσαν να συνδυαστούν από πολύ ικανά συστήματα προηγμένων κρατών ή υπηρεσιών. Για παράδειγμα το δίκτυο TOR το οποίο χρησιμοποιούμε στην μεθοδολογία μας, έχει κατηγορηθεί ότι περιλαμβάνει servers κρατικών υπηρεσιών, λειτουργώντας τελικά δηλαδή ως honeypot. Στον αντίποδα να αναφέρουμε επίσης ότι η διάρκεια εκτέλεσης του python script μας δεν διαρκεί περισσότερο από λίγα δευτερόλεπτα για να συγκεντρώσει τα δεδομένα από τις προγραμματιστικές γέφυρες που χρησιμοποιούμε εναντίον συγκεκριμένου χρήστη - στόχου, με αποτέλεσμα να αυξάνεται ο βαθμός δυσκολίας του εντοπισμού της επίθεσης.

3.1 Δημόσιο σημείο πρόσβασης στο Internet μέσω wi-fi και του Tails live usb

Ξεκινώντας την μεθοδολογία μας χρησιμοποιούμε ένα σημείο στο οποίο η πρόσβαση στο internet δεν μπορεί να συνδεθεί πίσω στην ταυτότητά μας (π.χ. το σπίτι μας, ή συγγενικό σπίτι, η εργασία μας κτλ). Βρισκόμαστε σε καφέ δημοφιλούς διεθνούς αλυσίδας στο κέντρο της Αθήνας, η οποία προσφέρει δωρεάν ασύρματη πρόσβαση στο Internet. Θα χρησιμοποιήσουμε τον φορητό μας υπολογιστή κάνοντας εκκίνηση από usb live διανομή Linux Tails. Η συγκεκριμένη διανομή προσφέρει ανωνυμία και προστασία της ιδιωτικότητας. Χρησιμοποιεί το δίκτυο Tor για να δρομολογήσει όλη την κίνηση στο Internet, με σκοπό να εξασφαλίσει την ανωνυμία. Η πλήρης ονομασία του Tor είναι: The Onion Router (Tor). Η λέξη onion (=κρεμμύδι) υποδηλώνει τα πολλαπλά "στρώματα" που χρησιμοποιεί κατά την λειτουργία του. Οι σκοποί που πλέον χρησιμοποιείται είναι πολλοί και δεν είναι όλοι κατ' ανάγκη ηθικοί ή νόμιμοι. Το δίκτυο Tor είναι ένα δίκτυο με συνδεδεμένους πολλούς υπολογιστές. Το πλήθος των χρηστών που είναι συνδεδεμένοι μέσω Tor, είναι σημαντικός παράγοντας καθώς όσο πιο μεγάλος ο αριθμός τους, τόσο πιο καλή η συνολική ποιότητα του δικτύου, μα και ακόμα πιο μεγάλη η ατομική ανωνυμία που μπορεί να επιτευχθεί. Θέλοντας λοιπόν να γίνει κάποια ανταλλαγή πληροφορίας από έναν υπολογιστή σε κάποιον άλλον, αυτή η πληροφορία, θα ταξιδέψει κρυπτογραφημένα, μέσα από ένα από τα χιλιάδες εναλλακτικά μονοπάτια που δίνουν οι συνδεδεμένοι υπολογιστές μεταξύ τους. Η σύνδεση σε κάποιον κόμβο είναι τυχαία κάθε φορά καθώς επίσης δεν είναι εμφανης ο τελικός προορισμός και η αφετηρία μέσα στο πακέτο ip, εξασφαλίζοντας την ανωνυμία. Ουσιαστικά μια ομάδα από relay nodes που βρίσκονται διασκορπισμένοι σε όλο τον κόσμο, χρησιμοποιούνται για την μείξη και κρυπτογράφηση της διαδικτυακής κίνησης σε πολλαπλά στρώματα (onion layers), έτσι ώστε στο τέλος της διαδικασίας να είναι εξαιρετικά δύσκολο να βρεθεί από που προήλθε εξαρχής το πακέτο πληροφορίας. Όταν ένας χρήστης Tor θέλει να αποκτήσει πρόσβαση σε μια διαδικτυακή τοποθεσία, ένα κρυπτογραφημένο αίτημα αποστέλλεται από τον browser του και περνά στο δίκτυο Tor. Ο πρώτος server που θα λάβει το request είναι ένας «guard» server, ο οποίος «ξεφλουδίζει» μέρος της κρυπτογράφησης και περνά το αίτημα σε ένα άλλο τυχαία επιλεγμένο διακομιστή. Η διαδικασία αυτή επαναλαμβάνεται μέχρι να αφαιρεθούν όλα τα στρώματα κρυπτογράφησης και ο τελευταίος server, ο κόμβος εξόδου, ή αλλιώς exit-node, προωθεί το αίτημα του προγράμματος περιήγησης του χρήστη στον πραγματικό server που φιλοξενεί τον επιλεγμένο ιστότοπο.

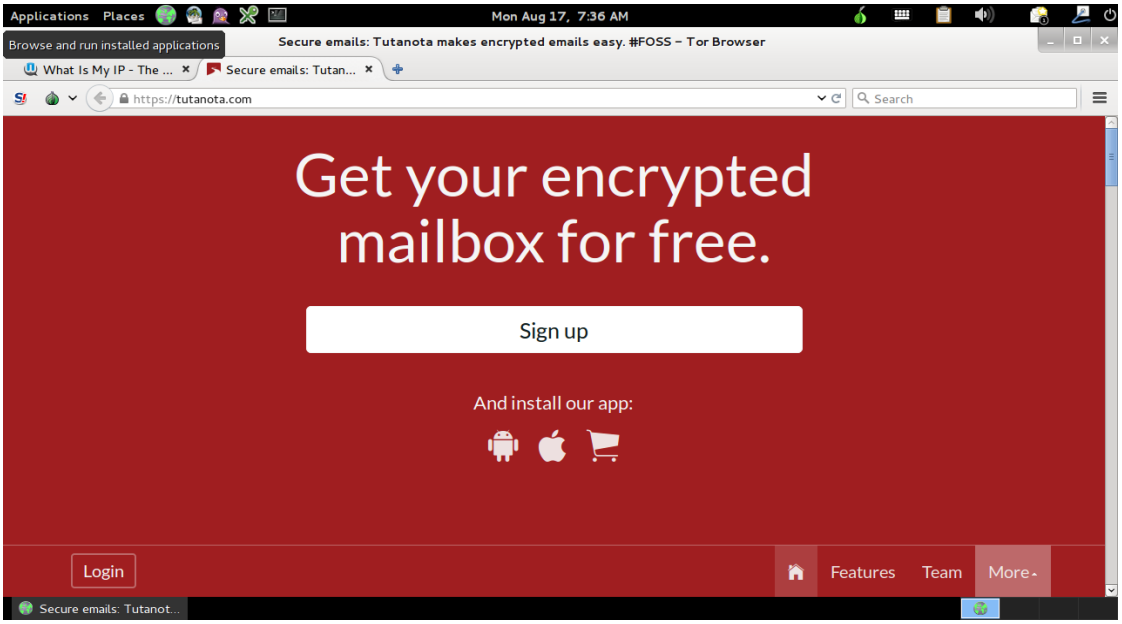


Εικόνα 1 *- Δρομολόγηση πακέτων μέσω του δικτύου TOR, για ανωνυμία.*

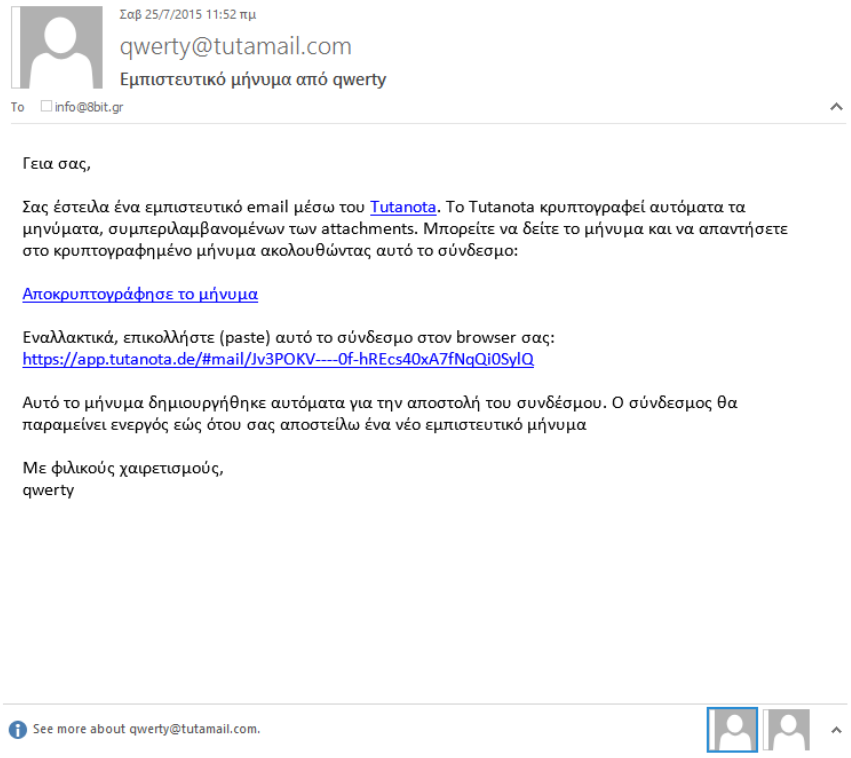
Εικόνα 2 *- Η διανομή Linux Tails αποκρύπτει την IP μας.*

3.2 Δημιουργία κρυπτογραφημένου mailbox

Με αρχικό σημείο αναφοράς το Tails live usb, δημιουργήσαμε ένα κρυπτογραφημένο mailbox μέσω της υπηρεσίας tutanota.com. Η συγκεκριμένη υπηρεσία κατά την εγγραφή δεν απαιτεί την δήλωση προσωπικών στοιχείων. Επιπλέον οι διευθύνσεις IP των εισερχόμενων και εξερχόμενων email δεν καταγράφονται στην συγκεκριμένη υπηρεσία, ενισχύοντας την ανωνυμία. Προσφέρεται επίσης κρυπτογράφηση των email (subject, content, attachment) και των επαφών. Η κρυπτογράφηση και αποκρυπτογράφηση των δεδομένων γίνεται στον τοπικό υπολογιστή και όχι σε κάποιο server (AES 128 bit και RSA 2048 bit). Το email μας είναι έτοιμο για χρήση μέσα σε λίγα λεπτά. Το email θα μας χρειαστεί ως ταυτότητα σε όλες τις υπόλοιπες Online υπηρεσίες που θα κάνουμε εγγραφή.



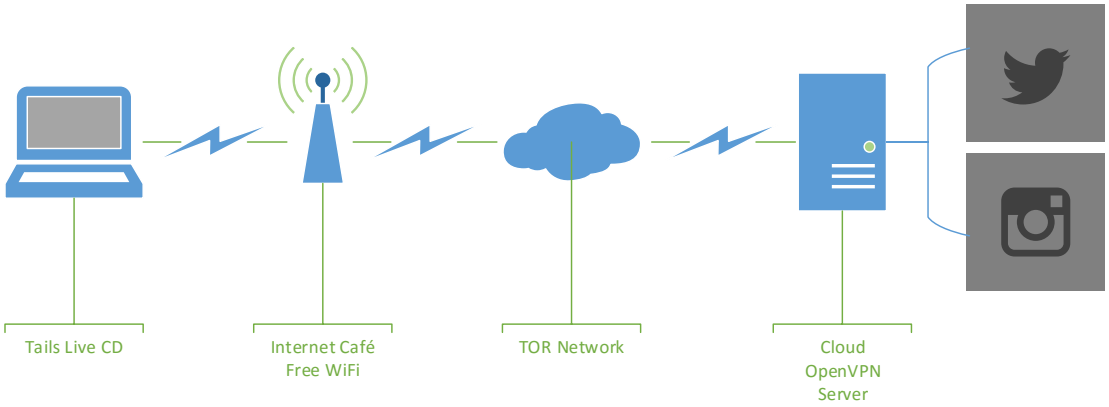
Εικόνα 3 *- Δημιουργία κρυπτογραφημένου mailbox.*



Εικόνα 4 *- Κρυπτογραφημένο email από το Tutamail.*

3.3 Αποκεντροποιημένο ηλεκτρονικό νόμισμα Bitcoin

Για να αυξήσουμε το ποσοστό της ανωνυμίας στην μεθοδολογία μας, θα χρησιμοποιήσουμε επιπρόσθετα ένα cloud server χαμηλών τεχνικών χαρακτηριστικών στον οποίο θα έχει εγκατασταθεί ένας vpn server σε Ubuntu 14.04 LTS 32bit. Με αυτό τον τρόπο θα προσθέσουμε στα εργαλεία μας άλλο ένα επίπεδο ανωνυμίας. Η σύνδεσή μας με το internet τότε θα γίνεται σύμφωνα με το παρακάτω σχήμα.



Εικόνα 5 *- Ανώνυμη διασύνδεση προς τους στόχους.*

Εδώ προκύπτει το επόμενο εμπόδιο στην προσπάθεια διασφάλισης της ανωνυμίας μας. Με ποιο τρόπο θα μπορέσουμε να πληρώσουμε την υπηρεσία του cloud server χωρίς να αποκαλύψουμε την πραγματική μας ταυτότητα. Η χρήση των πιστωτικών καρτών και δημοφιλών μεθόδων ηλεκτρονικών πληρωμών όπως το paypal αποκλείστηκαν καθότι απαιτούν την ταυτοποίηση μας. Ένας επιπλέον λόγος που μας απαγόρευε την χρήση πιστωτικών καρτών, κατά την περίοδο της συγγραφής της εργασίας αυτής, ήταν και ο περιορισμός τέτοιου είδους συναλλαγών για την Ελλάδα με αφετηρία τον Ιούλιο 2015 (capital controls). Η λύση στις συναλλαγές ήρθε με την χρήση του ηλεκτρονικού νομίσματος bitcoin και των παραγώγων, εναλλακτικών του.

3.4 Τι είναι το Bitcoin

Σύμφωνα με την σελίδα του, το bitcoin είναι «ένα συναινετικό δίκτυο που παρέχει τη δυνατότητα ενός νέου συστήματος πληρωμών και μιας εντελώς ψηφιακής μορφής χρημάτων. Είναι το πρώτο αποκεντρωμένο δίκτυο πληρωμής μεταξύ ομότιμων (peer-to-peer) που λειτουργεί από τους χρήστες του χωρίς κεντρική αρχή ή μεσάζοντες. Από τη σκοπιά του χρήστη, το Bitcoin είναι λίγο πολύ σαν τα μετρητά χρήματα του Διαδικτύου. Το Bitcoin [Bitcoin, 2015] μπορεί επίσης να θεωρηθεί ως το πιο περίφημο λογιστικό σύστημα τριπλής καταχώρησης που υπάρχει.»

Στην ερώτηση ποιος δημιούργησε το bitcoin πάλι σύμφωνα με την ίδια πηγή αναφέρονται: «Το Bitcoin είναι η πρώτη εφαρμογή μιας έννοιας που ονομάζεται «κρυπτονόμισμα», η οποία περιγράφηκε για πρώτη φορά το 1998 από τον Wei Dai στην λίστα αλληλογραφίας cypherpunks, υποστηρίζοντας την ιδέα μιας νέας μορφής χρήματος η οποία κάνει χρήση κρυπτογραφίας για να ελέγξει τη δημιουργία και τις συναλλαγές του, παρά μια κεντρική αρχή. Οι πρώτες προδιαγραφές του Bitcoin και η απόδειξη της έννοιας του δημοσιεύθηκαν το 2009 από τον Satoshi Nakamoto. Ο Satoshi αποσύρθηκε από το έργο αυτό στα τέλη του 2010 χωρίς να αποκαλύψει πολλά για τον εαυτό του. Από τότε, η κοινότητα του Bitcoin μεγάλωσε εκθετικά με πολλούς προγραμματιστές που ασχολούνται με το Bitcoin. Η "ανωνυμία" του Satoshi συχνά έθετε αδικαιολόγητες ανησυχίες, πολλές από τις οποίες συνδέονται με παρανοήσεις γύρω από την ανοιχτού κώδικα φύση του Bitcoin. Το λογισμικό και το πρωτόκολλο του Bitcoin εκδίδονται ανοιχτά προς όλους και οποιοσδήποτε προγραμματιστής μπορεί να επιθεωρήσει τον κώδικα ή να φτιάξει τις δικές του τροποποιημένες εκδόσεις του λογισμικού Bitcoin. Ακριβώς όπως και με τους τωρινούς προγραμματιστές, η επιρροή του Satoshi περιορίστηκε στις αλλαγές που έκανε και υιοθετήθηκαν από άλλους και συνεπώς δεν είχε τον έλεγχο του Bitcoin. Έτσι, η ταυτότητα του εφευρέτη του Bitcoin είναι πιθανόν τόσο σχετική σήμερα όσο και η ταυτότητα του προσώπου που εφηύρε το χαρτί.» [Bitcoin, 2015]

Τέλος στην ερώτηση πως λειτουργεί το bitcoin αναφέρονται οι παρακάτω πληροφορίες πάλι από την ίδια πηγή: «Από τη μεριά του χρήστη, το Bitcoin δεν είναι τίποτα περισσότερο από μια εφαρμογή κινητού τηλεφώνου ή υπολογιστή η οποία παρέχει ένα προσωπικό πορτοφόλι Bitcoin και επιτρέπει στο χρήστη να στέλνει και να λαμβάνει bitcoins μέσω αυτού . Έτσι λειτουργεί το Bitcoin για τους περισσότερους χρήστες. Στο παρασκήνιο, το δίκτυο Bitcoin μοιράζεται ένα δημόσιο λογιστικό βιβλίο (ledger) που ονομάζεται "block chain" (αλυσίδα των μπλοκ). Αυτό το βιβλίο περιέχει κάθε συναλλαγή που έχει ποτέ επεξεργαστεί από το δίκτυο, επιτρέποντας στον υπολογιστή του χρήστη να εξακριβώνει την εγκυρότητα της κάθε συναλλαγής. Η αυθεντικότητα της κάθε συναλλαγής προστατεύεται από ψηφιακές υπογραφές που αντιστοιχούν στις διευθύνσεις αποστολής, επιτρέποντας σε όλους τους χρήστες να έχουν πλήρη έλεγχο κατά την αποστολή bitcoins από τις δικές τους διευθύνσεις Bitcoin. Επιπλέον, ο καθένας μπορεί να επεξεργαστεί συναλλαγές που χρησιμοποιούν την υπολογιστική ισχύ

εξειδικευμένου υλισμικού (hardware) και να κερδίσει μια ανταμοιβή σε bitcoins για την υπηρεσία αυτή. Αυτό συχνά ονομάζεται εξόρυξη (mining).» [Bitcoin, 2015]

3.4.1 Πως αποκτάς bitcoin

Πρακτικά υπάρχουν τρεις τρόποι για να αποκτήσεις στην κατοχή σου bitcoin.

- Mining απευθείας σε bitcoin (cpu, gpu, asic mining).
- Αγορά σε bitcoin ATM με ευρώ.
- Mining σε εναλλακτικό ηλεκτρονικό νόμισμα και μετατροπή μέσω ανταλλακτηρίου σε bitcoin.

Η πρώτη μέθοδος δεν είναι πια ούτε αποδοτική, ούτε προσοδοφόρα σε μικρή κλίμακα. Αρχικά η εξόρυξη bitcoins γινόταν με εκτέλεση της εργασίας στην CPU, ενώ στην συνέχεια έγινε εισαγωγή της έννοιας του GPU mining, δηλαδή της εκτέλεσης της ίδιας εργασίας αλλά με την χρήση της δύναμης της κάρτας γραφικών AMD και NVIDIA. Η μέθοδος αυτή μπορούσε να γίνει scale up, δηλαδή να τρέξουμε το πρόγραμμα mining σε πολλαπλές cpu / cores και σε πολλαπλές κάρτες γραφικών στο ίδιο σύστημα ή να γίνει scale out, δηλαδή να χρησιμοποιήσουμε πολλαπλούς υπολογιστές που δεν είναι αναγκαίο να έχουν την ίδια σύνθεση ή επεξεργαστική ισχύ, για να τρέξουμε συνεργατικά και παράλληλα το πρόγραμμα mining. Στην πορεία εμφανίστηκαν στην αγορά ειδικά επεξεργαστικά κυκλώματα ASIC [Barr, 2007:123] (Application-specific integrated circuit) για την εκτέλεση bitcoin mining software με σημαντικά πλεονεκτήματα (ταχύτητα, ευκολία χρήσης, φυσικές διαστάσεις, ενεργειακή κατανάλωση) σε σύγκριση με το CPU/GPU mining. Την εποχή συγγραφής της εργασίας αυτής όπως αναφέραμε, δεν είναι οικονομικά συμφέρουσα η συγκεκριμένη μέθοδος.

Η δεύτερη μέθοδος που είναι νέα σχετικά στο συγκεκριμένο πεδίο για την Ελλάδα, αφορά την χρήση αυτόματων μηχανών τραπεζικών συναλλαγών, οι οποίες όμως έχουν ως μέσο συναλλαγής bitcoins και ευρώ. Πρακτικά μπορείς να αλλάξεις ευρώ σε bitcoins. Στην Αθήνα τον Αύγουστο / Σεπτέμβριο 2015, λειτουργούν 2 τέτοια bitcoin ATM, ενώ έχει ανακοινωθεί η εγκατάσταση νέων.

Η τρίτη μέθοδος, η οποία είναι και αυτή που επιλέχθηκε στην συγκεκριμένη μεθοδολογία, είναι η παραγωγή εναλλακτικών cryptocoins με την χρήση cpu και gpu και στην συνέχεια η ανταλλαγή τους σε bitcoins σε online ανταλλακτήρια ψηφιακών νομισμάτων. Θα πρέπει να επισημάνουμε ότι ακόμα και ο συγκεκριμένος τρόπος δεν είναι οικονομικά προσοδοφόρος καθότι η αξία του ρεύματος που απαιτείται για την ανακάλυψη των cryptocoins είναι μεγαλύτερη από την αξία των cryptocoins. Παρόλα αυτά μας προσφέρει το αποκεντροποιημένο, μη ανιχνεύσιμο σύστημα πληρωμών το οποίο θα χρησιμοποιήσουμε για να αγοράσουμε online υπηρεσίες που θα μας βοηθήσουν στο σκοπό μας.

3.4.2 Monero (XMR)

Το ψηφιακό νόμισμα που επιλέχτηκε για να κάνουμε mining, είναι το Monero (XMR). Πρόκειται για ένα ασφαλές, ιδιωτικό, μη ανιχνεύσιμο ψηφιακό νόμισμα. Παράλληλα αποτελεί προϊόν open source. Προσφέρει ασφάλεια συναλλαγών, χρησιμοποιώντας ένα κατανεμημένο συναινετικό δίκτυο (P2P), όπου κάθε συναλλαγή είναι κρυπτογραφημένη. Είναι ασφαλές καθότι οι συναλλαγές δεν είναι δημόσια ορατές στο καθολικό αρχείο συναλλαγών (blockchain). Τέλος είναι μη ανιχνεύσιμο καθότι με την αξιοποίηση ring signatures στην κρυπτογραφία, είναι αδύνατο να εντοπιστούν οι συναλλαγές και να συνδεθούν σε συγκεκριμένο χρήστη.

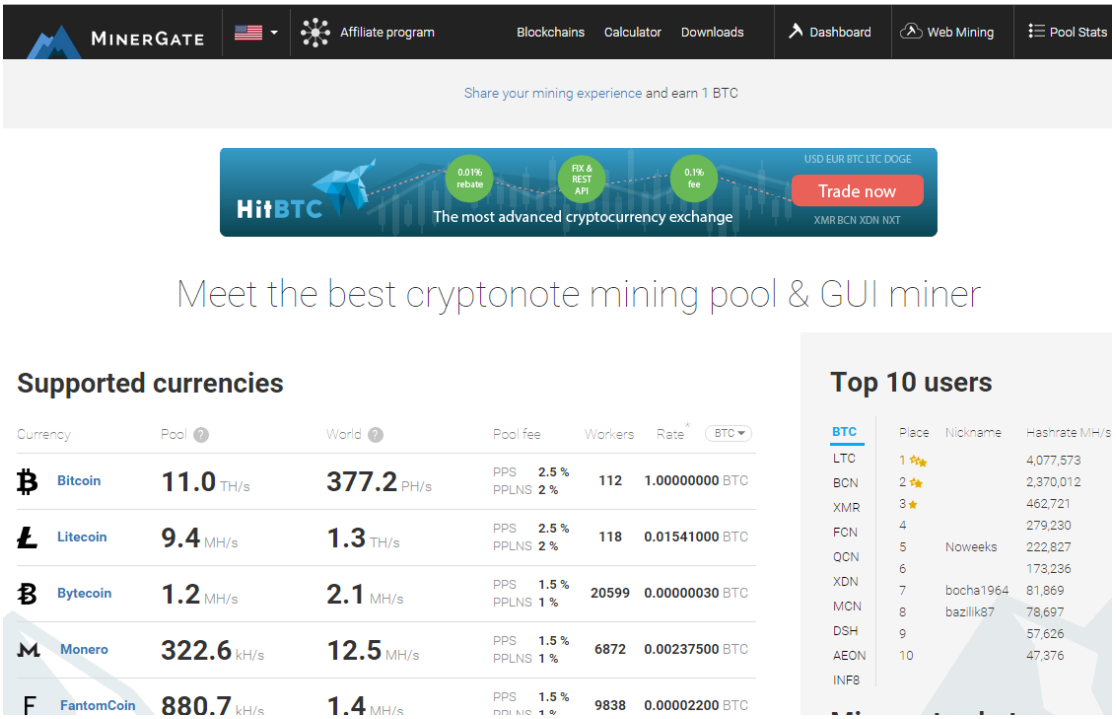
Για την ανακάλυψη (mining) Monero cryptocoins θα αξιοποιήσουμε δύο υπολογιστές υψηλών τεχνικών χαρακτηριστικών, οι οποίοι θα εργάζονται αποκλειστικά για τον σκοπό αυτό συνεχώς. Συνοπτικά τα τεχνικά χαρακτηριστικά καθενός από τους δυο υπολογιστές είναι τα εξής:

- CPU: Intel i7-4790k (4 cores – 8 threads) με υπερχρονισμό στα 4,5 GHz
- RAM: 16 GB DDR3
- GPU: AMD Radeon 280X

3.4.3 Mining Pools

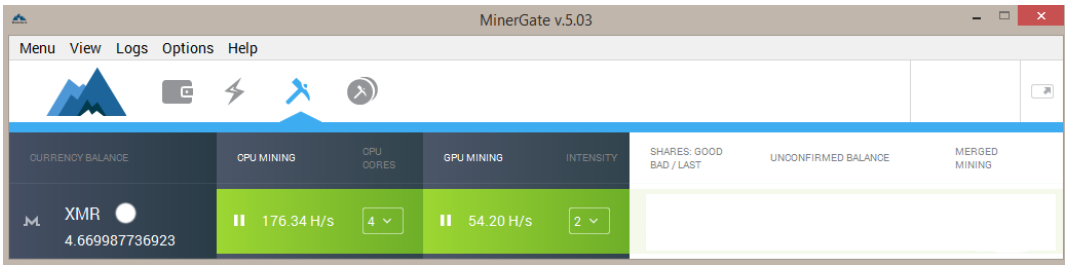
Ακόμα και με την χρήση πολλών και ισχυρών συστημάτων είναι πολύ δύσκολο να ανακαλυφθούν νέα ψηφιακά νομίσματα. Για να μην αποκλείονται και οι κάτοχοι χαμηλότερων δυνατοτήτων υπολογιστικών μηχανών από την διαδικασία του mining, έχουν δημιουργηθεί online mining pools, κοινόχρηστων δεξαμενών που συλλογικά δουλεύουν προς την ανακάλυψη νέων cryptocoins. Ο κάθε ένας από τους συμμετέχοντες στην όλη διαδικασία, ανταμείβεται με ένα ποσοστό από το νέο cryptocoin.

Υπάρχουν πολλές διαφορετικές πλατφόρμες mining pools. Από την έρευνα που διεξάχθηκε και με κριτήρια την ευκολία mining και το cryptocoin που επιλέχθηκε, δηλαδή το monero, επιλέχθηκε το MinerGate. Η εγγραφή έγινε μέσω του tutamail email που ενεργοποιήσαμε νωρίτερα και πάντα μέσα από το tor network που προσφέρει το Tails live usb.

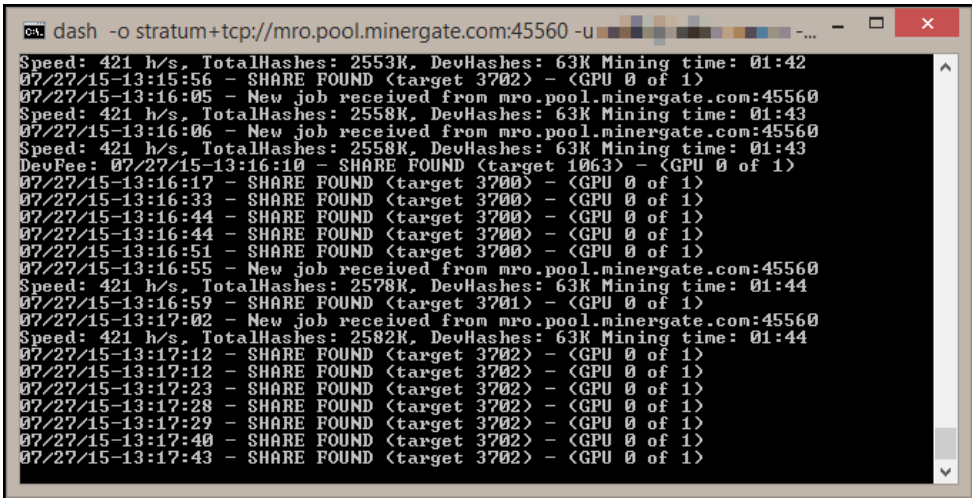


Εικόνα 6 *- To mining pool minergate.*

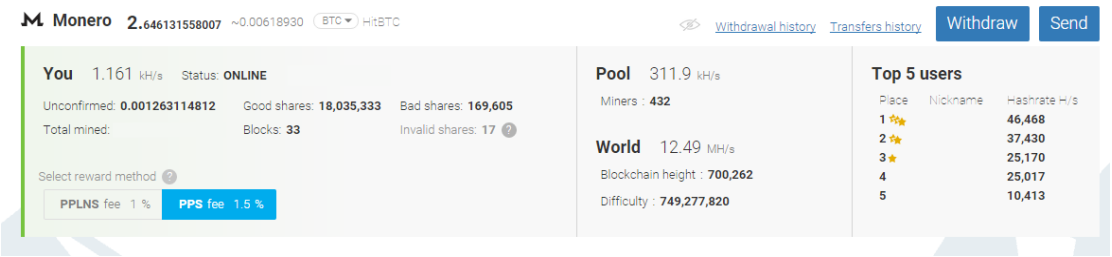
Ένα άλλο σημαντικό πλεονέκτημα του Minergate είναι ότι προσφέρει τον δικό του gui ή command line client για cpu - gpu mining σε Windows και Ubuntu Linux λειτουργικά συστήματα.



Εικόνα 7 *- Minergate windows client.*



Εικόνα 8 *- Minergate terminal client.*



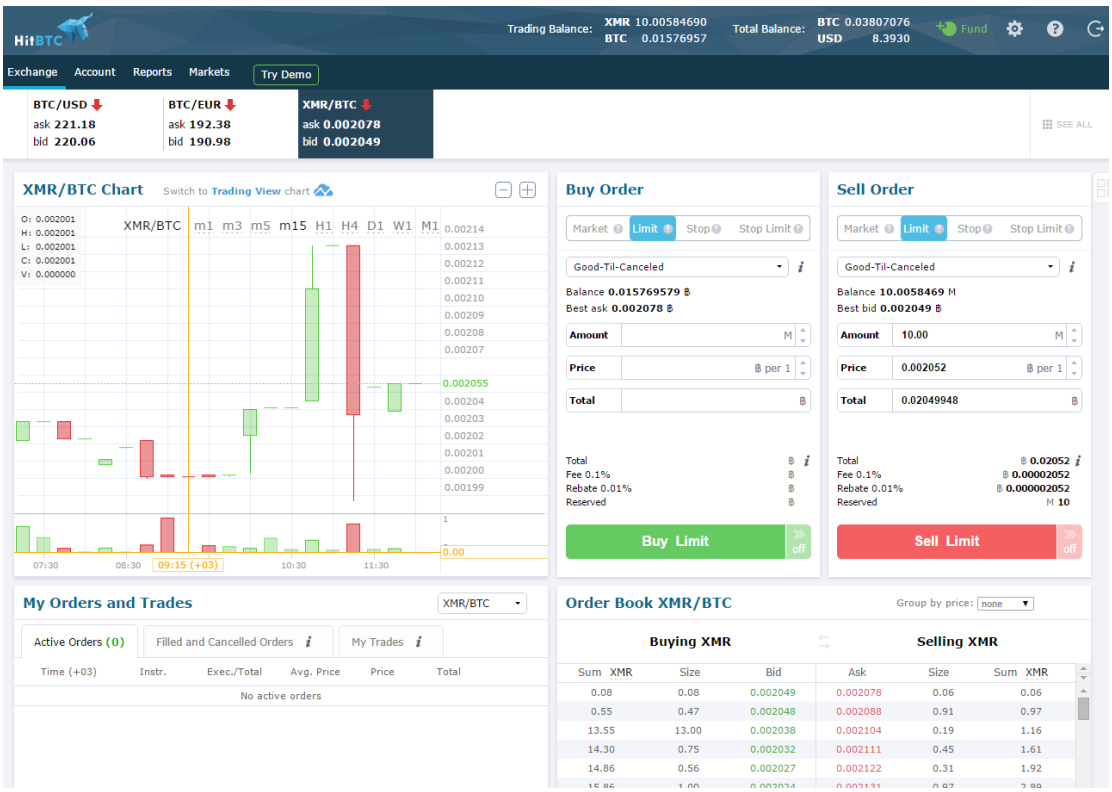
Εικόνα 9 *- Minergate online dashboard.*

Με συνεχή χρήση των δυο παραπάνω συστημάτων και μετά από βελτιστοποίηση software και hardware καταφέραμε να επιτύχουμε ταχύτητα γύρω 1500 kH/s [Paar, 2009:120], το οποίο μεταφράζεται σε περίπου 10 monero cryptocoins την εβδομάδα ή

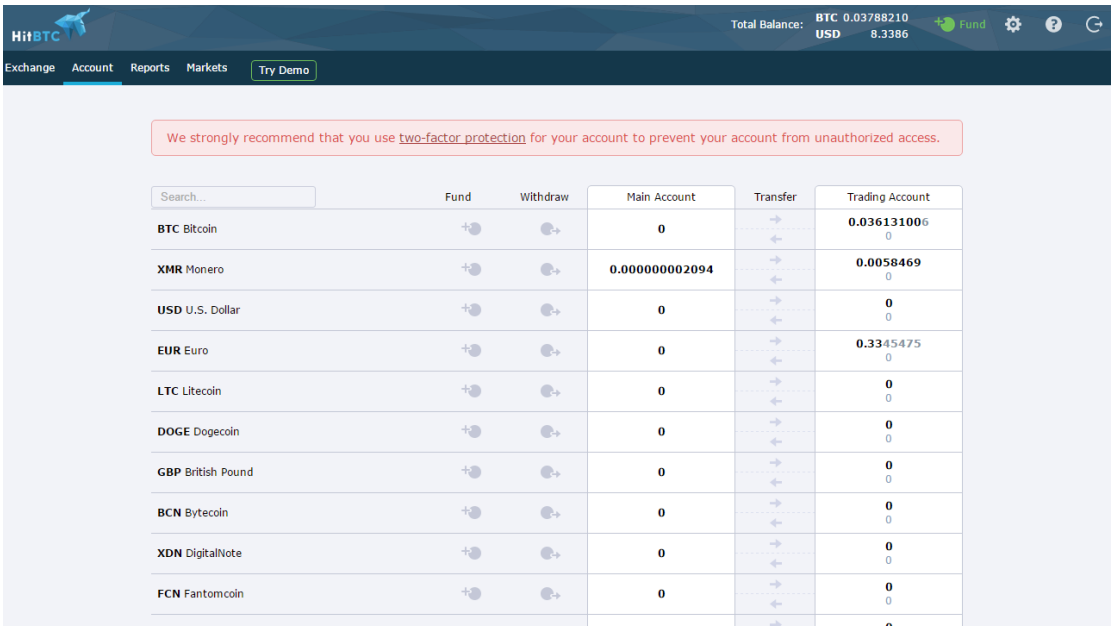
περίπου 0,032 bitcoins ή 5-5,5€ αναλόγως με την τρέχουσα ισοτιμία. Το ποσό αυτό είναι επαρκές για την κάλυψη του μηνιαίου κόστους ενός μικρού μεγέθους cloud server στην συνέχεια.

3.4.4 Virtual Currency Exchange

Τα monero cryptocoins δεν είναι ιδιαίτερα δημοφιλή ψηφιακά νομίσματα, σε αντίθεση με τα bitcoins. Αυτό σημαίνει ότι δεν μπορούμε να τα χρησιμοποιήσουμε απευθείας για την αγορά του cloud server μας. Θα πρέπει να τα μετατρέψουμε σε bitcoins έτσι ώστε να μπορέσουμε να αγοράσουμε online cloud servers για την συνέχεια. Τα monero cryptocoins συγκεντρώνονται στην πλατφόρμα του minergate στον λογαριασμό μας. Για να τα ανταλλάξουμε σε bitcoins θα πρέπει να τα μεταφέρουμε σε ένα ανταλλακτήριο ηλεκτρονικών νομισμάτων. Επιλέξαμε το hitbtc.com καθότι συνεργάζεται ενδογενώς με το minergate.

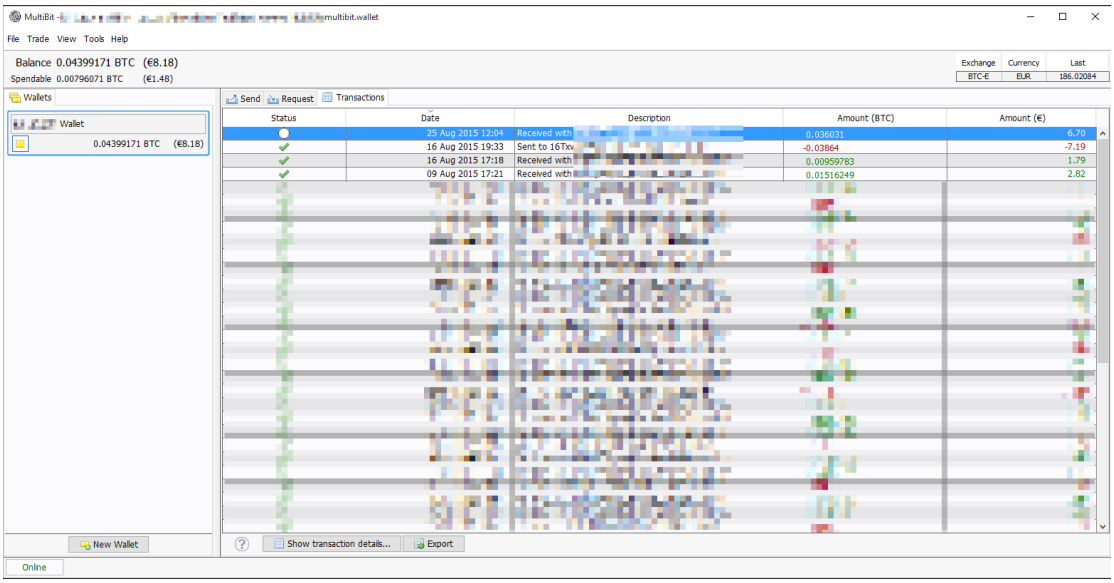


Εικόνα 10 *- HitBTC ανταλλακτήριο ψηφιακών νομισμάτων.*



Εικόνα 11 *- HitBTC ανταλλακτήριο ψηφιακών νομισμάτων.*

Μετά την μεταφορά των Monero cryptocoins μας στο hitbtc και την ανταλλαγή τους σε bitcoins, έχουμε την δυνατότητα είτε να αποστείλουμε τα bitcoin μας σε προσωπικό μας ηλεκτρονικό πορτοφόλι στον υπολογιστή μας (CryptoCoin Wallet) ή να τα χρησιμοποιήσουμε μέσα από το HitBTC για να αγοράσουμε το cloud server μας.

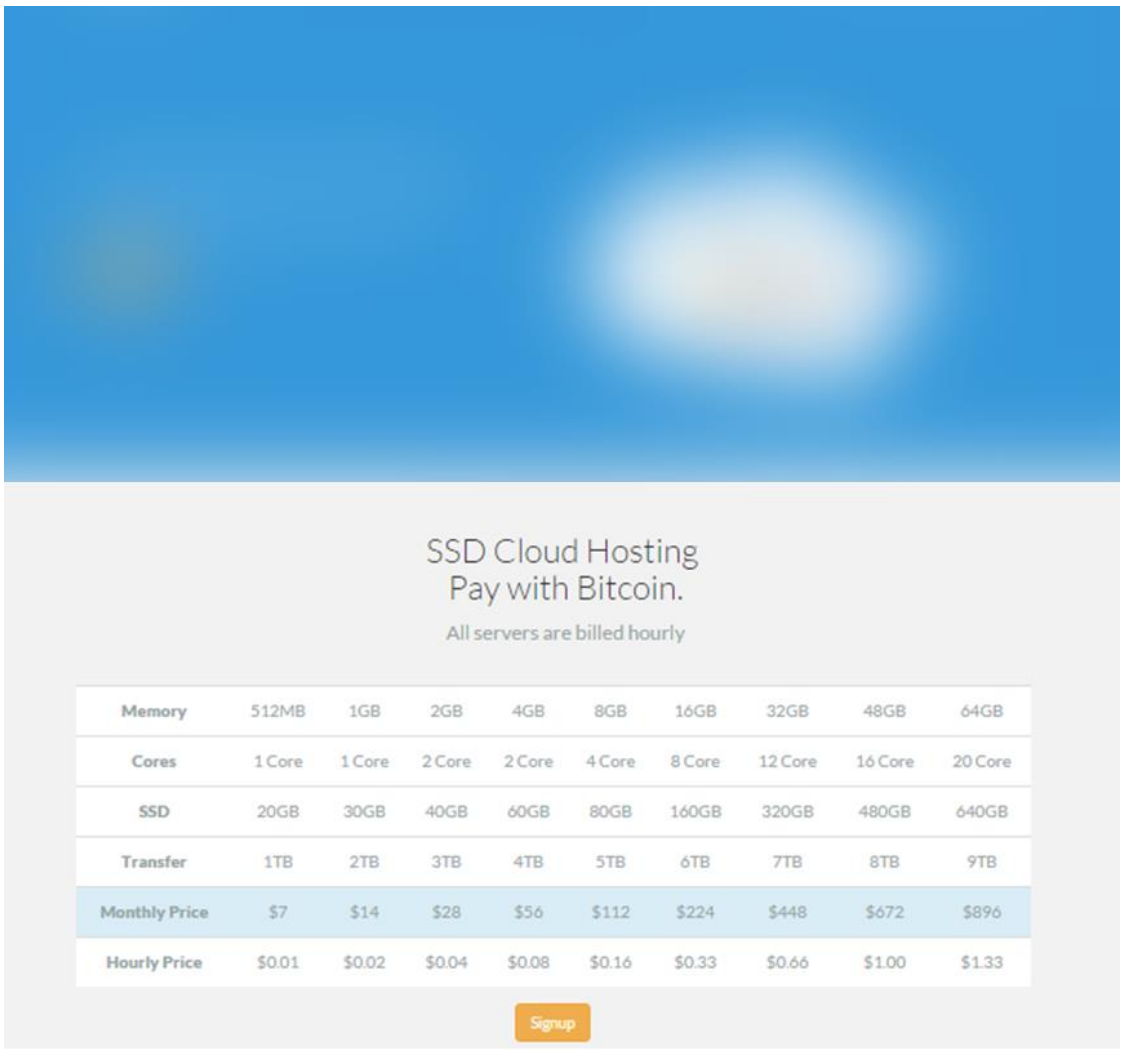


Εικόνα 12 *- MultiBit cryptocoin wallet.*

3.5 Cloud Servers

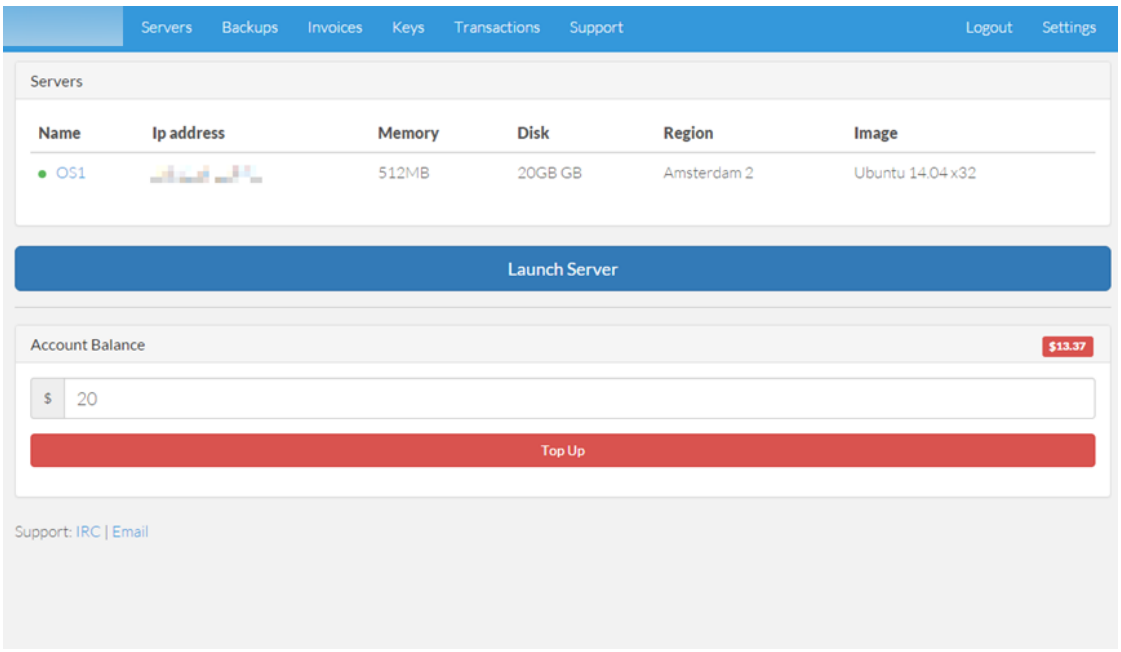
Με τον όρο cloud server ορίζεται ένα virtual machine το οποίο βρίσκεται στην υποδομή ενός παρόχου IaaS (Infrastructure as a Service) και είναι προσβάσιμο από εξωτερικούς πελάτες μέσω συγκεκριμένης αποκλειστικής διεύθυνσης IPv4. Οι υπηρεσίες που μπορούν να προσφέρει ποικίλουν, καθώς και οι προδιαγραφές του σε CPU, RAM, storage, bandwidth, λειτουργικό σύστημα. Στην αγορά υπάρχουν μεγάλες εταιρίες που προσφέρουν αντίστοιχες υπηρεσίες όπως η Microsoft, Google, Amazon, VMWare, Rackspace, Digital Ocean κ.α. που όμως απαιτούν ταυτοποίηση του πελάτη και αγορά μέσω πιστωτικής κάρτας.

Για να αποφύγουμε την ταυτοποίηση, θα χρησιμοποιήσουμε υπηρεσίες cloud hosting που δέχονται ως πληρωμή bitcoins τα οποία συλλέξαμε νωρίτερα. Κατά την περίοδο συγγραφής της παρούσας διπλωματικής εργασίας ανακαλύψαμε δύο τέτοιες υπηρεσίες και κατόπιν σύγκρισης των δύο λύσεων καταλήξαμε στην συγκεκριμένη η οποία βασίζεται στην υποδομή της DigitalOcean που είναι υψηλά στην προτίμηση των προγραμματιστών. Απαιτεί 7$ πληρωμένα σε bitcoins για κάθε μήνα λειτουργίας του cloud server που επιλέξαμε (1 CPU Core, 512MB RAM, 20GB SSD storage, 1TB Data transfer).

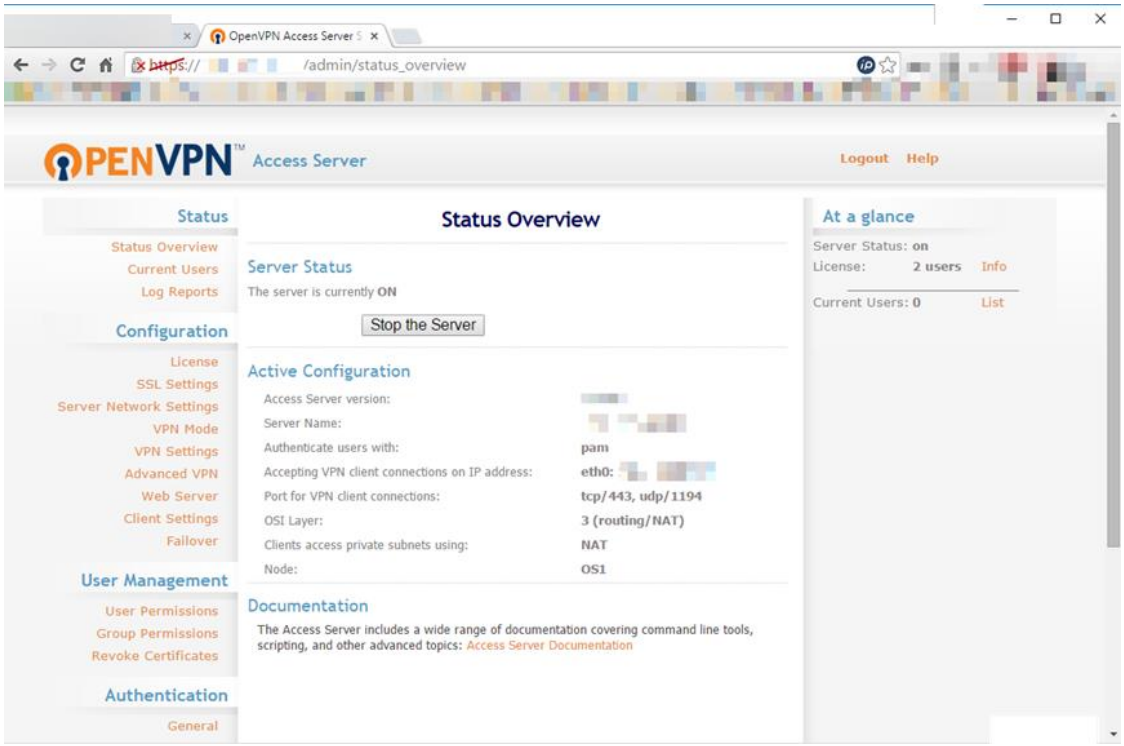


Εικόνα 13*- Cloud Hosting μέσω πληρωμής με bitcoins.*

Μετά την μεταφορά των bitcoins είχαμε πολύ γρήγορα διαθέσιμο τον cloud server μας με λειτουργικό σύστημα Ubuntu 14.04 LTS server x32. Μετά τις απαραίτητες ρυθμίσεις και την διαδικασία hardening του συστήματος μας, εγκαταστήσαμε το OpenVPN service. Η πρόσβαση στον vpn server θα γίνεται μέσω του δικτύου TOR, έτσι ώστε να έχουμε διπλασιάσει τους τρόπους απόκρυψης της πραγματικής μας ταυτότητας και ταυτόχρονα τα ίχνη μας να έχουν γίνει ακόμα περισσότερο μη ανιχνεύσιμα.



Εικόνα 14 *- O cloud server.*



Εικόνα 15 *- OpenVPN server.*

Στο σημείο αυτό έχουμε επιτύχει να δημιουργήσουμε επιτυχώς και ανώνυμα την υποδομή που παρουσιάζεται στην Εικόνα 4.

Κεφάλαιο 4
Twitter API

4.1 Εισαγωγή στο Twitter

Το Twitter είναι μια πλατφόρμα κοινωνικής δικτύωσης η οποία επικεντρώνεται στην γρήγορη και σύντομη επικοινωνία μεταξύ των χρηστών της, μέσω μηνυμάτων των 140 χαρακτήρων. Η ταχύτητα και η ευκολία χρήσης του, το έχουν αναδείξει σε κορυφαίο επικοινωνιακό μέσο της σύγχρονης εποχής. Σύμφωνα με το ίδιο το Twitter και με ημερομηνία αναφοράς την 30η Σεπτεμβρίου 2015, κάθε μήνα εξυπηρετεί 320 εκατομμύρια χρήστες, με 1 δισεκατομμύριο επισκέψεις, ενώ οι χρήστες smartphone αντιπροσωπεύουν το 80% του συνόλου [Twitter, 2015]. Απασχολεί 4300 υπαλλήλους σε 35 τοποθεσίες ανά τον κόσμο. Οι χρήστες του Twitter το 2013 δημιουργούσαν μέσα στην πλατφόρμα περισσότερα από 400 εκατομμύρια μηνύματα κάθε μέρα [WashingtonPost, 2013].

Ο όγκος της συγκεκριμένης πληροφορίας που συσσωρεύεται στο Twitter, είναι διαθέσιμος στον προγραμματιστή μέσω δύο API, που διαφοροποιούνται όσον αφορά τον σχεδιασμό και τον τρόπο πρόσβασης:

- REST API [Twitter, 2013] που είναι βασισμένα στην αρχιτεκτονική REST [Ong, 2015:210], όπου το πρόγραμμα ανασύρει μια συγκεκριμένη πληροφορία από το σύστημα (pull for data retrieval).

- Streaming API [Twitter, 2014] , όπου προσφέρεται μια συνεχόμενη ροή πληροφορίας από το Twitter στο πρόγραμμα (push for data retrieval). Το πρόγραμμα ζητά στην έναρξη της επικοινωνίας μιας συγκεκριμένης ροής δεδομένων από το Twitter και η ροή των δεδομένων ενημερώνεται συνεχώς και αδιαλείπτως. Λόγω της φύσης του χρησιμοποιείται στην παρακολούθηση σε πραγματικό χρόνο των μηνυμάτων του Twitter.

Στην εφαρμογή μας χρησιμοποιούμε το REST API για να αποκτήσουμε προγραμματιστική πρόσβαση στην πλατφόρμα του Twitter. Το REST API πιστοποιεί τις εφαρμογές μέσω του ασφαλούς πρωτοκόλλου ταυτοποίησης OATH [Twitter, 2014] και λαμβάνει τις απαντήσεις σε μορφή JSON.

4.1.1 OAuth

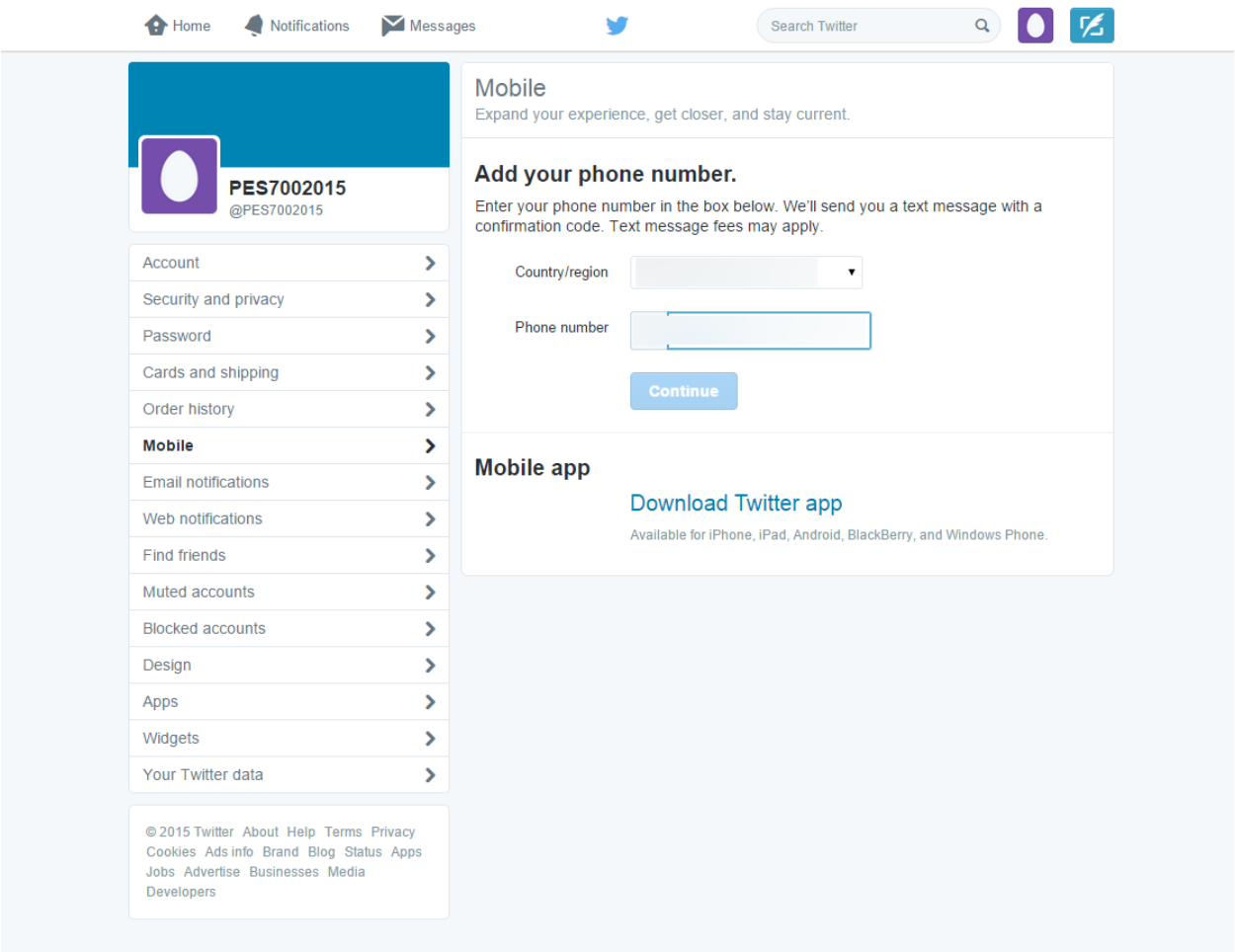
Το Open Authentication (OAuth) είναι ένα ανοικτό πρότυπο για πιστοποίηση και χρησιμοποιείται από το Twitter για να δώσει πρόσβαση σε προστατευμένη πληροφορία. Σε σχέση με την πρακτική της αυθεντικοποίησης μέσω ονόματος χρήστη και κωδικό, το OAuth προσφέρει μια ασφαλέστερη εναλλακτική χρησιμοποιώντας την λογική του three way handshake. Επιπλέον ο χρήστης είναι σίγουρος ότι δεν μοιράζεται με την εφαρμογή τον συνδυασμό όνομα χρήστη / κωδικού. Η διαδικασία της δημιουργίας μιας εφαρμογής για το Twitter απαιτεί την αντίστοιχη δημιουργία τεσσάρων διαπιστευτηρίων:

- **Consumer key** (γραμμή 11). Οι εφαρμογές (consumers) εγγράφονται στο Twitter μέσω ενός μοναδικού αριθμού.
- **Consumer secret** (γραμμή 12). Οι εφαρμογές (consumers) αφού εγγραφούν και λάβουν consumer_key, αποκτούν και τον αντίστοιχο κωδικό πρόσβασης στο API.
- **Access token** (γραμμή 13). Αφορούν τον μοναδικό αριθμό που αντιστοιχεί στον χρήστη / προγραμματιστή του Twitter. Δεν μπορεί να αλλάξει.
- **Access token secret** (γραμμή 14). Αφορά τον μοναδικό κωδικό που αντιστοιχεί στον χρήστη / προγραμματιστή του Twitter. Δεν μπορεί να αλλάξει.

4.2 Δημιουργία εφαρμογής στο Twitter

4.2.1 Εισαγωγή

Η εγγραφή μιας νέας εφαρμογής στο Twitter είναι εύκολη υπόθεση, αλλά προϋποθέτει την ολοκλήρωση ενός απαραίτητου βήματος, η οποία είναι η επιβεβαίωση της ταυτότητάς μας μέσω SMS σε κινητό τηλέφωνο, όπως φαίνεται στην παρακάτω εικόνα.



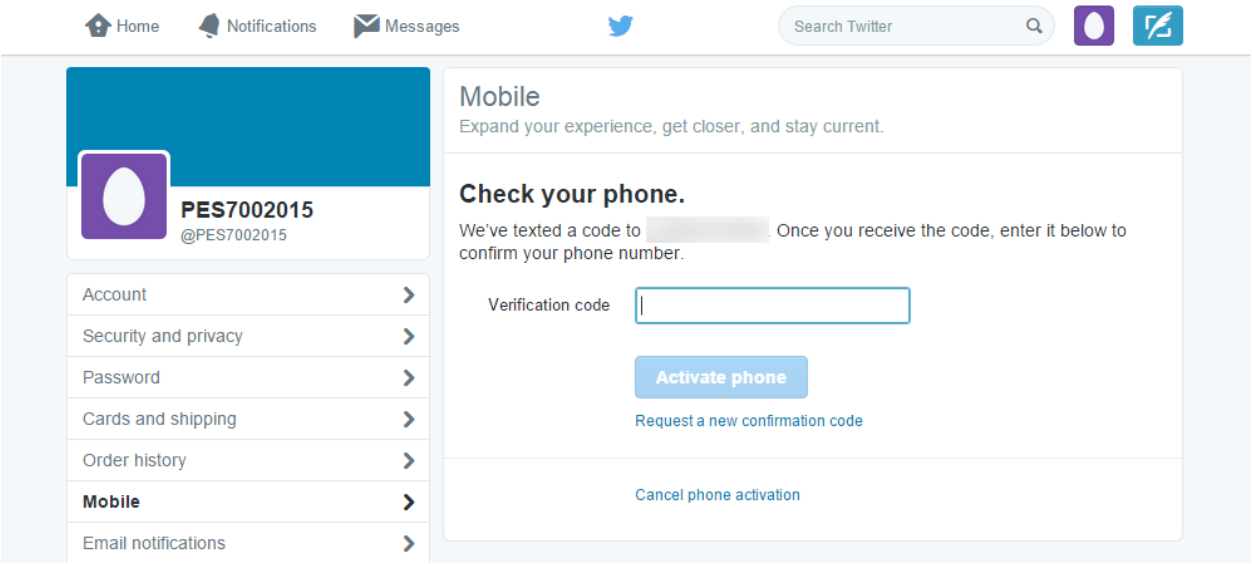
Εικόνα 16 *- Πιστοποίηση χρήστη μέσω SMS σε κινητό τηλέφωνο για την δημιουργία Twitter app.*

Το συγκεκριμένο βήμα είναι απαραίτητο και δεν μπορεί να παρακαμφθεί. Σύμφωνα με την μεθοδολογία μας και για τους ερευνητικούς σκοπούς της συγκεκριμένης εργασίας, ενεργούμε πάνω στην λογική του OSINT και του κακόβουλου χρήστη που δεν πρέπει να αποκαλύπτει σε κανένα βήμα την πραγματική ταυτότητα του. Για την Ελλάδα και σύμφωνα με τον Ν. 3783/2009 "Ταυτοποίηση των κατόχων και χρηστών εξοπλισμού και υπηρεσιών κινητής τηλεφωνίας και άλλες διατάξεις" (ΦΕΚ 136/Α/7-8-09) [ΕΕΤΤ, 2009] , οι κάτοχοι και χρήστες εξοπλισμού και υπηρεσιών κινητής τηλεφωνίας προπληρωμένου χρόνου ομιλίας, συνδρομητών με συμβόλαιο, ή άλλης μορφής κινητής τηλεπικοινωνίας υποχρεούνται να ταυτοποιηθούν. Η τελική ημερομηνία για την

ολοκλήρωση της διαδικασίας ήταν η 30η Ιουλίου 2010. Το συγκεκριμένο γεγονός απέκλεισε την χρήση ελληνικού αριθμού κινητής τηλεφωνίας για την διαδικασία της επιβεβαίωσης.

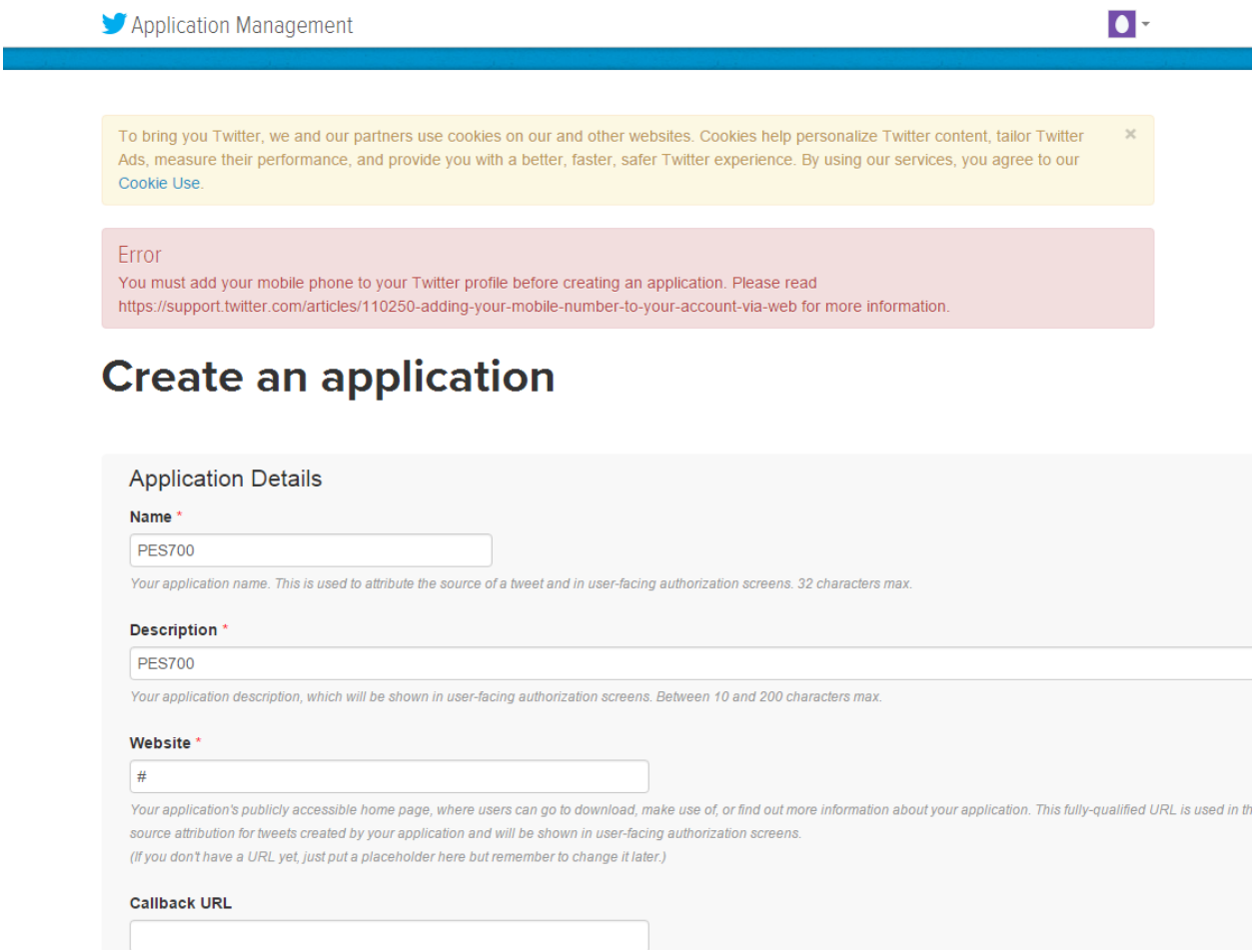
Στην συνέχεια στραφήκαμε σε διαδικτυακές υπηρεσίες που προσφέρουν εικονικούς αριθμούς κινητής τηλεφωνίας για διάφορες χώρες του κόσμου. Δυστυχώς και σε αυτή την περίπτωση και μετά από πολλές αποτυχημένες προσπάθειες, δεν καταφέραμε να ολοκληρώσουμε την διαδικασία. Το Twitter χρησιμοποιεί ειδικούς αλγόριθμους και τεχνικές για να αποκλείσει τηλεφωνικούς αριθμούς που έχουν χρησιμοποιηθεί σε διαφορετικές περιπτώσεις (black lists).

Η επόμενη προσπάθεια μας αφορούσε την διερεύνηση του νομικού καθεστώτος ταυτοποίησης χρηστών κινητής τηλεφωνίας σε γειτονικές χώρες της Ελλάδας. Για την Βουλγαρία και για τον Αύγουστο του 2015 ισχύει καθεστώτος ταυτοποίησης, όπως και στην Ελλάδα. Η λύση τελικά δόθηκε με την χρήση προπληρωμένου καρτοκινητού με χώρα προέλευσης την Κύπρο (Αύγουστος 2015), όπου για την ενεργοποίηση δεν απαιτείται ταυτοποίηση. Με αυτόν τον τρόπο ολοκληρώθηκε η διαδικασία ταυτοποίησης στο Twitter, χωρίς να έχουμε αποκαλύψει πραγματικά στοιχεία μας, όπως φαίνεται στην παρακάτω εικόνα.

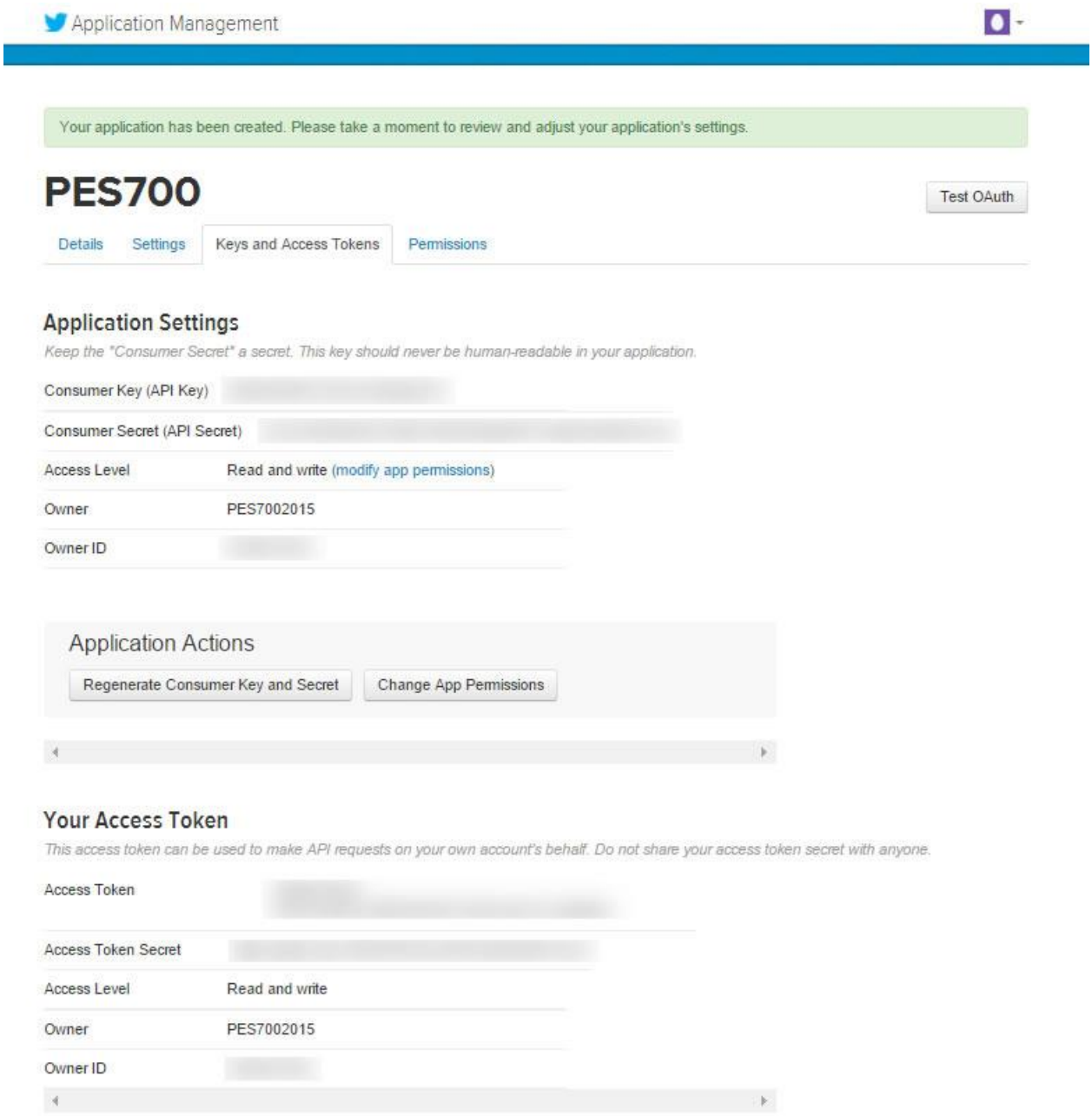


Εικόνα 17 *- Διαδικασία ταυτοποίησης χρήστη Twitter Dev.*

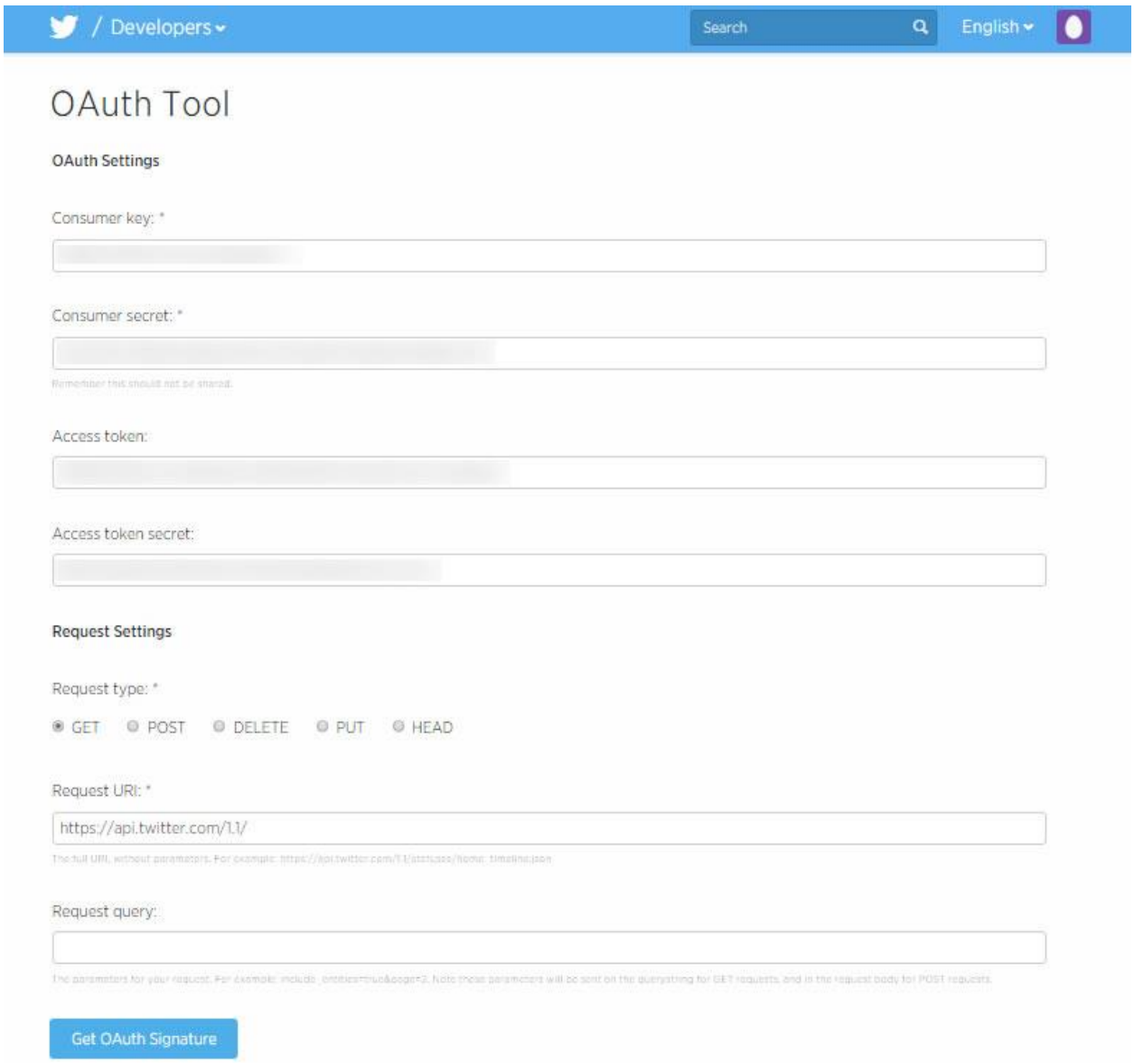
Αφού ξεπεράστηκε το εμπόδιο της ταυτοποίησης μέσω κινητού για το Twitter, μπορούμε να δημιουργήσουμε την εφαρμογή μας και να λάβουμε τα απαραίτητα OAuth κλειδιά για το SocialMap python script.



Εικόνα 18 *- Εγγραφή εφαρμογής στο Twitter.*



Εικόνα 19 *- Consumer key, secret και Access token, secret για την χρήση στο SocialMap python script.*



Εικόνα 20 *- Twitter OAuth tool.*

4.2.2 Ανάκτηση των πρόσφατων μηνυμάτων ενός χρήστη του Twitter με την χρήση του REST API

Με την χρήση της μεθόδου GET user_timeline [Twitter, 2014], μπορούμε να συλλέξουμε τις τελευταίες δημοσιεύσεις ενός χρήστη του Twitter. Με την συγκεκριμένη μέθοδο μπορούμε να λάβουμε τις τελευταίες 3200 δημοσιεύσεις του χρήστη στόχου. Η απάντηση από το Twitter έχει την μορφή JSON, ενώ υπάρχει ο περιορισμός των 300 αιτήσεων ανά

χρονικό παράθυρο 15 λεπτών από την εφαρμογή μας. Η συγκεκριμένη εντολή χρησιμοποιείται μέσα από τις βιβλιοθήκες της Python στην γραμμή 243 του script μας. Για να ικανοποιήσουμε τους περιορισμούς του REST API διαβάζουμε σε ομάδες των 200 tweets από το timeline του στόχου μας. Ο μέγιστος συνολικός αριθμός έχει οριστεί σε 1000 tweets.

4.2.3 Στρατηγικές για τον προσδιορισμό της θέσης προέλευσης του μηνύματος

Η πληροφορία της γεωγραφικής θέσης στο Twitter είναι διαθέσιμη σε δύο διαφορετικές περιοχές:

- Πληροφορία coordinates: Οι χρήστες κατ' επιλογή μπορούν να φανερώνουν την θέση που βρίσκονται μέσω του GPS δέκτη του smartphone τους, την στιγμή που δημιουργούν το μήνυμα του Twitter. Η πληροφορία αυτή περιέχει μεγάλη ακρίβεια. Ακολουθεί την μορφή του geoJSON [Geojson, 2013] (longitude, latitude). Ένα παράδειγμα φαίνεται στην παρακάτω εικόνα:

```
1  "coordinates":
2  {
3      "coordinates":
4      [
5          -75.14310264,
6          40.05701649
7      ],
8      "type":"Point"
9  }
```

Εικόνα 21 *- Παράδειγμα geoJSON από το Twiiter REST API.*

- Πληροφορία Location: Πρόκειται για πεδίο αποθηκευμένο από τον χρήστη, που αναφέρει ονομαστικά την πόλη, περιοχή. Δεν παρέχει μεγάλη ακρίβεια. Για να χρησιμοποιηθεί θα πρέπει να μετατραπεί σε γεωγραφικές συντεταγμένες, πράγμα

το οποίο το πετυχαίνουμε με την χρήση του Google maps geocoding API. Ένα παράδειγμα φαίνεται στην παρακάτω εικόνα:

```
1  "location":"San Francisco, CA"
```

Εικόνα 22 *- Παράδειγμα πεδίου location από το Twitter REST API.*

Κεφάλαιο 5
Instagram API

5.1 Εισαγωγή στο Instagram

Η εφαρμογή δημιουργήθηκε από δύο απόφοιτους του Πανεπιστημίου του Στάντφορντ, τους Κέβιν Σίστρομ και Μάικ Κρίγκερ και ξεκίνησε τον Οκτώβριο του 2010. Μέσα σε δύο μήνες μόνο, ο αριθμός των εγγεγραμμένων χρηστών έφτασε το 1.000.000. Σήμερα η εφαρμογή μετράει 20 δισεκατομμύρια φωτογραφίες από όλο τον κόσμο και 200 εκατομμύρια ενεργούς χρήστες. Το 2012 η εφαρμογή αγοράστηκε από το Facebook, ενώ από τον Ιούνιο του 2013 προστέθηκε και η δυνατότητα εγγραφής και διαμοιρασμού video [Instagram, 2015].

Μερικά από τα χαρακτηριστικά του Instagram είναι τα ακόλουθα:

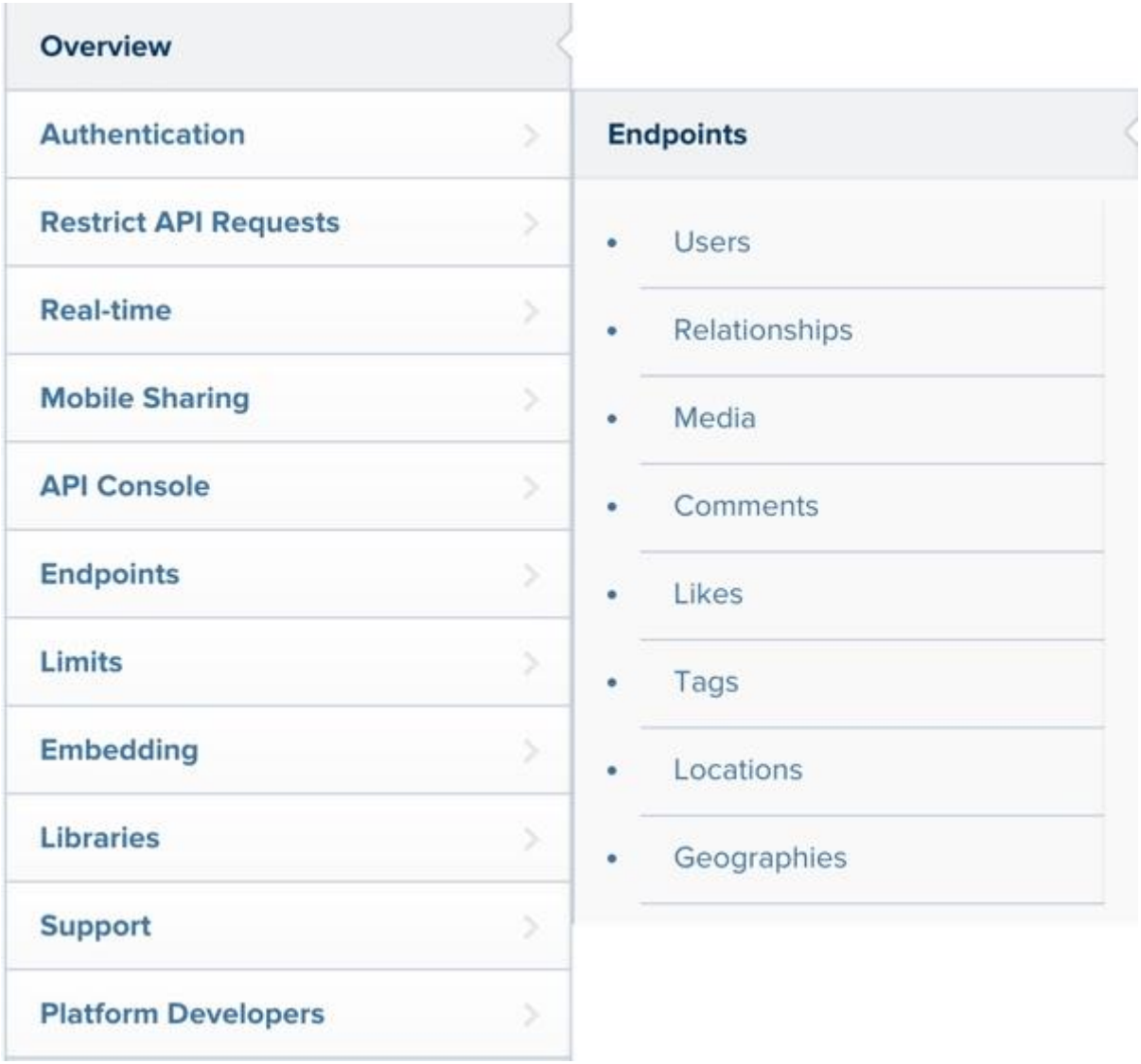
- Ο χρήστης μπορεί να τραβήξει μια φωτογραφία, video ή να χρησιμοποιήσει κάποιο που είναι ήδη στην κατοχή του.
- να εφαρμόσει φίλτρα και διάφορα εφέ.
- να κάνει αναφορά και άλλους χρήστες στις δημοσιεύσεις του (tag).
- να προσθέσει τοποθεσία και περιγραφή και στην συνέχεια να τα κοινοποιήσει κατευθείαν στο Facebook, στο Twitter, στο Flickr και άλλα δίκτυα κοινωνικής δικτύωσης.

5.2 Instagram API

Ο όγκος των δεδομένων που υπάρχει στο Twitter, είναι διαθέσιμος στον προγραμματιστή μέσω Instagram API:

Instagram API (basic endpoints)

- GET /locations/location-id.
  Πληροφορίες που σχετίζονται με την γεωγραφική τοποθεσία.
- GET /locations/location-id/media/recent
  Δοσμένης συγκεκριμένης γεωγραφικής τοποθεσίας, επέστρεψε τα σχετικά media (φωτογραφίες, videos).
- GET /locations/search
  Αναζήτηση με βάση γεωγραφικές συντεταγμένες μιας τοποθεσίας.



Εικόνα 23 *- Instagram API Endpoints*

Οι παρακάτω παράμετροι μπορούν να προσφέρουν διάφορες πληροφορίες σχετικά με τους χρήστες του Instagram [Instagram, 2015].

- **Χρήστες**

  Η παράμετρος χρήστες επιτρέπει να γίνει αναζήτηση χρηστών σύμφωνα με το όνομα, ενώ μπορεί να γίνει αναζήτηση σύμφωνα με τα άτομα που ακολουθούν οι χρήστες αλλά και τα μηνύματα που ανταλλάσσουν.
- **Σχέσεις**

Η παράμετρος σχέσεων επιτρέπει την άντληση των πληροφοριών σχετικά με την λίστα ακολούθων καθώς επίσης τις συνδέσεις που δημιουργούνται από τις σχέσεις.

- **Media**

  Η παράμετρος των media μπορεί να αντλήσει πληροφορίες σχετικά με το video ή την φωτογραφία που αναρτάται στο Instagram, δίνοντας την δυνατότητα την εξαγωγή πληροφοριών σχετικά με το μέσο όπου παράχθηκε το video ή φωτογραφία καθώς και η τοποθεσία όπου έγινε.

- **Σχόλια, Ευχαριστίες, Ετικέτες**

  Η παράμετρος των σχολίων, ευχαριστιών και των ετικετών απαριθμεί των αριθμό των σχολίων των σχολίων, ευχαριστιών αλλά και των ετικετών που έχουν γίνει σε κάποιο μέσο (φωτογραφία ή video).

- **Τοποθεσίες**

  Τα μέσα κοινωνικής δικτύωσης έχουν το κάθε ένα από αυτά την δικιά τους βάση δεδομένων αναφορικά με την τοποθεσία. Το Instagram την δικιά του βάση την ονομάζει Locations όπου δίνει το δικαίωμα σε όποιον χρησιμοποιεί το Api της εξάγει πληροφορίες σχετικά με τις συντεταγμένες όπου δημοσιευθήκαν τα video ή φωτογραφίες.

Η πιστοποίηση των χρηστών γίνεται όπως και στην περίπτωση του Twitter μέσω του πρωτόκολλου OAuth 2.0, που αναλύσαμε στο προηγούμενο κεφάλαιο.

5.3 Instagram Real-Time API

Με το Instagram Real-Time API [Instagram, 2014], μπορεί κάποιος να ελέγξει σε πραγματικό χρόνο την δραστηριότητα των χρηστών αναφορικά με τους χρήστες, τις ετικέτες αλλά και τοποθεσίες των media.

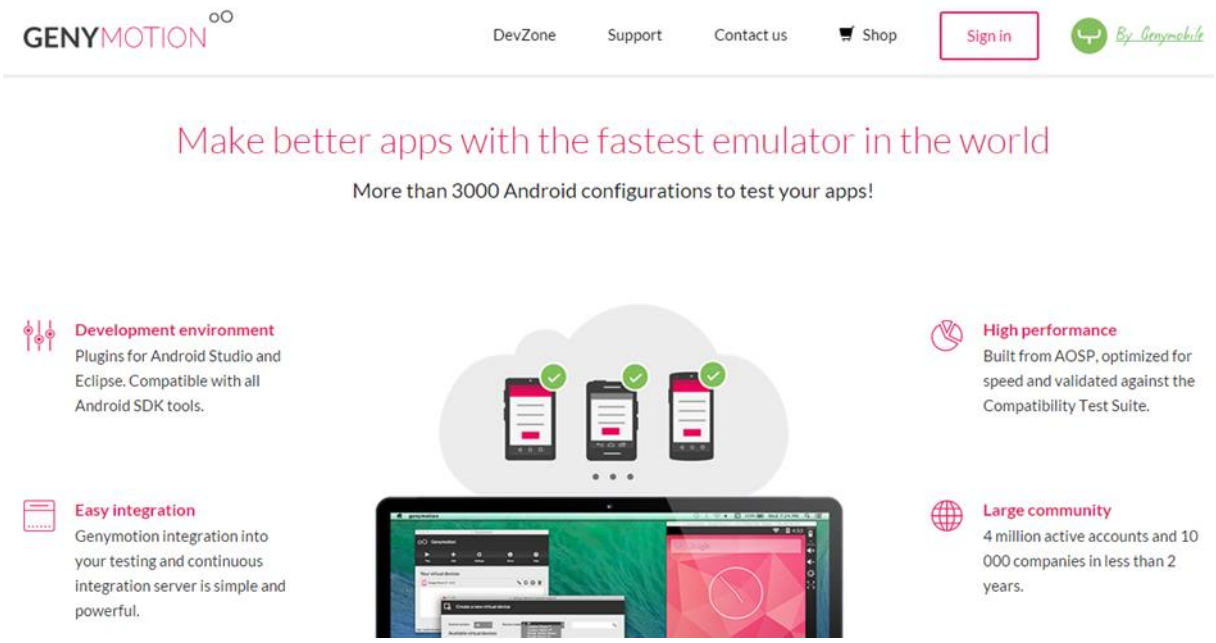
Έτσι όταν γίνεται αναζήτηση του χρήστη μέσω του Instagram Real-Time API τότε εξάγονται πληροφορίες σχετικά με τις νέες του δημοσιεύσεις. Όταν γίνεται αναζήτηση σε σχέση με τις ετικέτες τότε γίνεται εξαγωγή πληροφορίας ανάλογα με την ετικέτα που

δημοσιεύεται στο εκάστοτε media. Μέσα από την αναζήτηση τοποθεσιών λαμβάνονται ειδοποιήσεις για νέες φωτογραφίες ή video που έχουν αναρτηθεί και η ετικέτα με μια συγκεκριμένη τοποθεσία. Επίσης μπορεί να γίνει εξαγωγή δεδομένων αναφορικά με νέες φωτογραφίες που έχουν αναρτηθεί σε μια αυθαίρετη γεωγραφική θέση.

5.4 Εγγραφή στο Instagram API

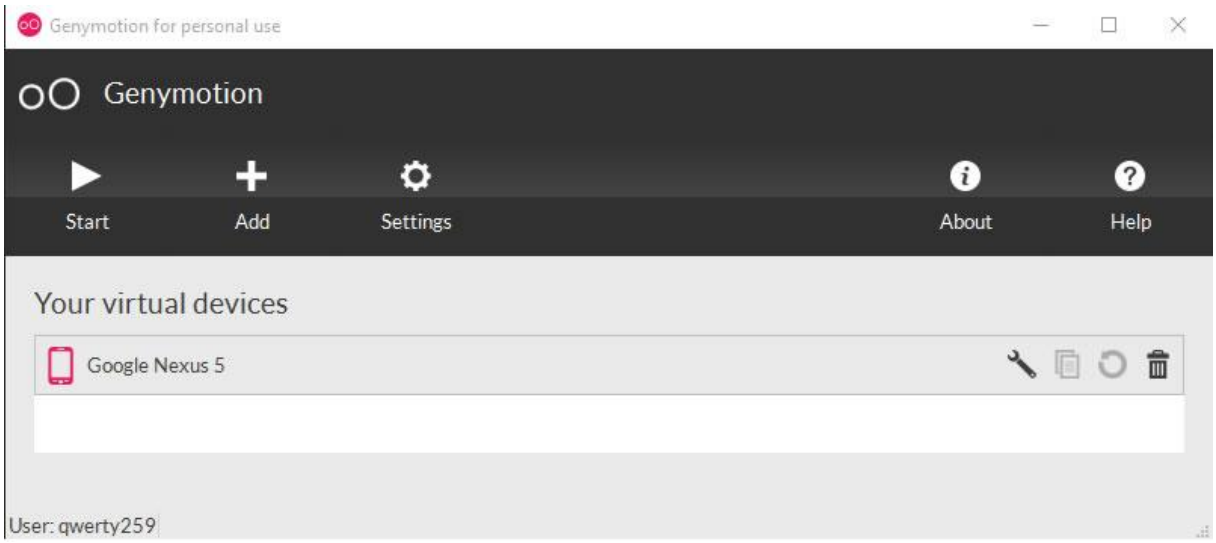
Η διαδικασία εγγραφής στο Instagram διαφέρει από αυτή του Twitter καθότι απαιτεί την χρήση Android ή Apple smartphone, για την ολοκλήρωσή της. Δεν μπορεί να γίνει μέσω του website της εφαρμογής, καθότι είναι πρωτίστως μια εφαρμογή που απευθύνεται σε χρήστες smartphones. Η ιδιαιτερότητα αυτή, θα μπορούσε να αποτελέσει πρόβλημα στην μεθοδολογία μας που προσεγγίζεται από την σκοπιά του κακόβουλου χρήστη (hacker). Δεν θα μπορούσαμε να χρησιμοποιήσουμε μια κανονική συσκευή android καθότι αυτό θα εξέθεται στοιχεία ταυτότητας, όπως το μοναδικό αριθμό IMEI [European, 2009] της συσκευής μας ή το Wi-Fi access point στο οποίο έχουμε συνδεθεί, το οποίο μπορεί να οδηγήσει σε συμπέρασμα γεωγραφικής τοποθεσίας.

Η ενδεδειγμένη λύση στο παραπάνω πρόβλημα δόθηκε μέσω ενός προγράμματος android smartphone emulator και συγκεκριμένα του Genymotion [Genymotion, 2014]. Μια από τις εκδόσεις του, προορίζεται για προσωπική χρήση και είναι δωρεάν.



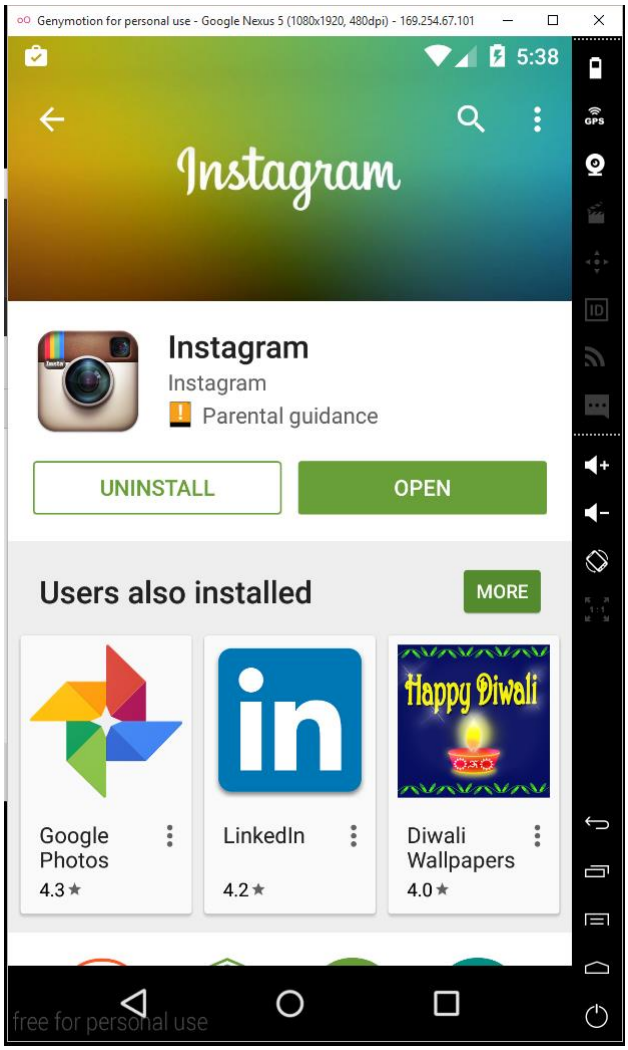
Εικόνα 24 *- Genymotion Android Emulator*

Μετά την εγκατάσταση του προγράμματος, δημιουργήσαμε μια εικονική συσκευή Google Nexus 5 και στην συνέχεια έγινε εγκατάσταση του Google Play Store app.



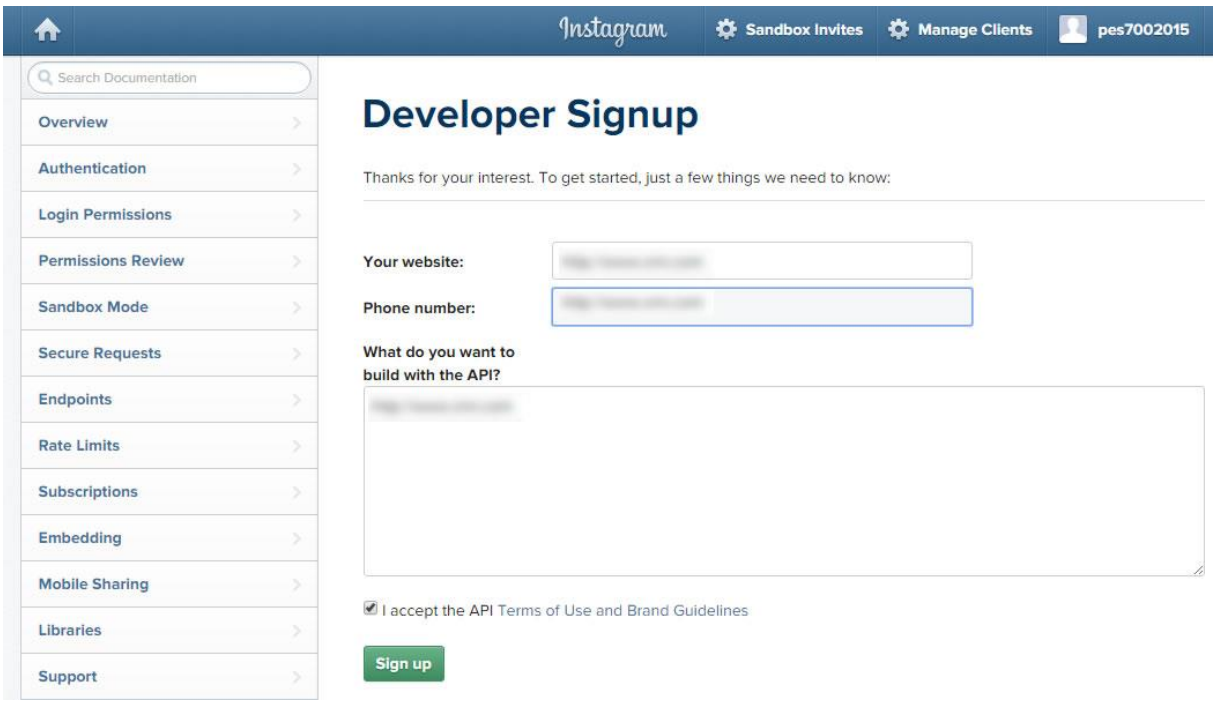
Εικόνα 25 *- Google Nexus 5 Virtual Machine*

Στην συνέχεια εγκαταστήσαμε το Instagram app από το Google Play Store και εγγραφήκαμε μέσω του λογαριασμού email στο tutanota.com.



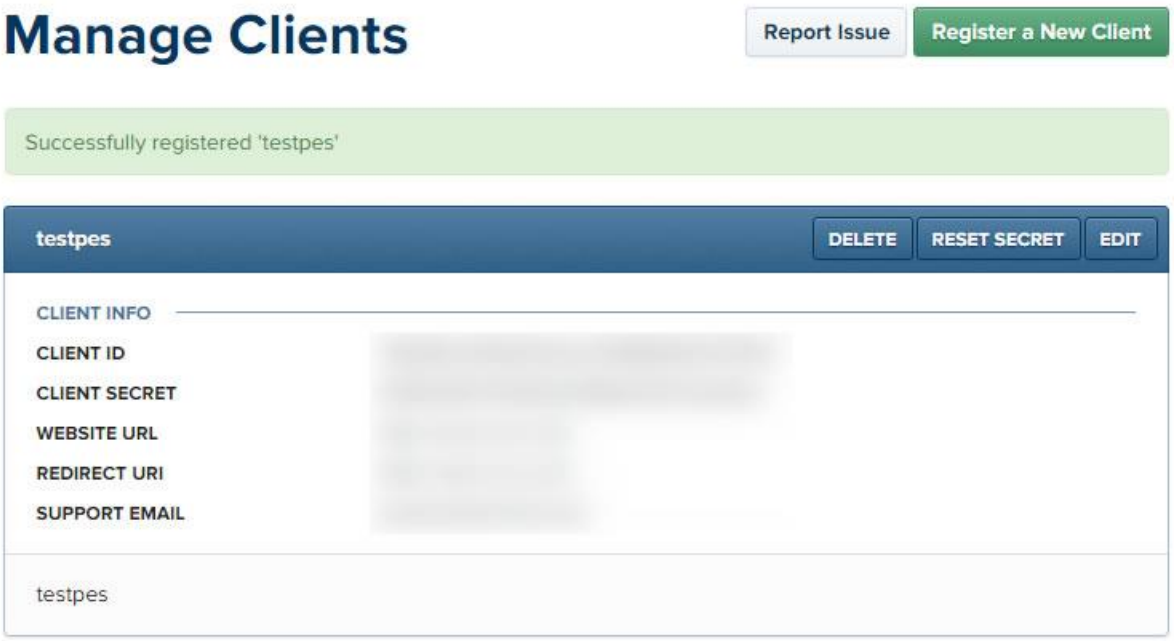
Εικόνα 26 *- Εγκατάσταση Instagram app στην εικονική συσκευή Google Nexus 5.*

Έχοντας ολοκληρώσει με επιτυχία την εγγραφή στην εφαρμογή του Instagram μπορούμε να κάνουμε τώρα χρήση του API, μέσα από το website για τους προγραμματιστές. Στην εγγραφή της εφαρμογής μας θα ζητηθεί ένας αριθμός κινητού τηλεφώνου, αντίστοιχα με την περίπτωση του Twitter API. Χρησιμοποιούμε την ίδια μέθοδο απόκρυψης της ταυτότητας μας, όπως και στο Twitter API.



Εικόνα 27 *- Εγγραφή στο Instagram API*

Με την επιτυχή εγγραφή μας της εφαρμογής μας, έχουμε στην διάθεσή μας τα μοναδικά κλειδιά Client ID και Client Secret (γραμμές 299 και 300 στο SocialMap python script).



Εικόνα 28 *- Κλειδιά εφαρμογής Instagram API.*

Κεφάλαιο 6
Google API

Στην σημερινή εποχή, η χρήση χαρτών σε μια εφαρμογή που περιλαμβάνει γεωγραφικές συντεταγμένες είναι συνηθισμένη. Υπάρχουν διάφορες υπηρεσίες που προσφέρουν λύσεις που αφορούν απεικόνιση σε χάρτη, όπως οι Yahoo maps, Bing maps, με πιο διαδεδομένη την Google maps που τελικά επιλέχθηκε. Τα διαφορετικά API μας βοηθούν έτσι ώστε να απεικονίσουμε δεδομένα από άλλες εφαρμογές πάνω σε ένα Google maps χάρτη. Στην περίπτωσή μας αφορούν συντεταγμένες που έχουν προκύψει από την δραστηριότητα του χρήστη – στόχου από τις δύο πλατφόρμες κοινωνικής δικτύωσης Twitter και Instagram.

Στο SocialMap python script, γίνεται χρήση δύο Google API και συγκεκριμένα του Google maps javascript [Google, 2014] και του Google maps geocoding [Google, 2014]. Στις γραμμές 21 (google geocode api key) και 78 (google maps api key) του Python script μας, θα πρέπει να εισάγουμε τα κλειδιά από τα αντίστοιχα API. Στο συγκεκριμένο κεφάλαιο θα αναλύσουμε τις δυνατότητες, τους περιορισμούς τους και την χρήση τους στο πρόγραμμά μας.

6.1 Google maps JavaScript API

Το Google maps API παρέχει τη λειτουργικότητα των google maps με γρήγορο και άμεσο τρόπο. Είναι μια web εφαρμογή υπηρεσιών, η οποία παρέχεται από την Google και προσφέρει οδικούς χάρτες και υπηρεσίες πλοήγησης σε δικτυακούς τόπους ή σε mobile εφαρμογές.

Οι υπηρεσίες που προσφέρονται είναι οι εξής:

- Δημιουργία και εμφάνιση χάρτη.
- Εισαγωγή markers (pinpoints), πολυγώνων, popups, polylines, Info Window.
- Event handlers.
- Geocoder: Υπηρεσία μετατροπής συντεταγμένων/διευθύνσεων.
- Direction: Σχεδίαση δρομολογίου και οδηγιών. Πλοήγηση διαδρομής (με αυτοκίνητο, με δημόσια μέσα μεταφοράς ή με τα πόδια).
- εντοπισμός επιχειρήσεων σε χώρες σε όλο τον κόσμο.

Το Google Maps API είναι δωρεάν για εμπορική χρήση, υπό τον όρο ότι η εφαρμογή στην οποία χρησιμοποιείται θα είναι προσβάσιμη στο κοινό χωρίς να χρεώνει τον χρήστη του για κάθε πρόσβαση, και δεν παράγουν περισσότερα από 25.000 προσβάσεις χάρτη ανά ημέρα [Google, 2014]. Φυσικά υπάρχουν premium πακέτα επί πληρωμή τα οποία καλύπτουν αυξημένες ανάγκες σε αιτήσεις και χρήση. Για την περίπτωση του SocialMap python script, οι προδιαγραφές του δωρεάν πακέτου είναι υπεραρκετές.

Web	STANDARD	PREMIUM
Google Maps JavaScript API Google Static Maps API Google Street View Image API	Free until exceeding 25,000 map loads per day for 90 consecutive days $0.50 USD / 1,000 additional map loads above 25,000 per day after reaching 25,000 map load / 90 day usage limit, up to 1,000,000 daily	Pricing based on volume required Premium features: • Guaranteed ad-free • Image size up to 2048 x 2048 pixels
Google Maps Embed API	Unlimited free usage	—

Εικόνα 29 *- Οι χρεώσεις του Google maps javascript API.*

Η ερώτηση στο Google Maps API γίνεται με την αποστολή ενός αιτήματος HTTP GET στην Web εφαρμογή, και επιστρέφει την απάντηση σε μορφή XML ή JSON. Η εφαρμογή η οποία περιγράφεται σε αυτό το έγγραφο χρησιμοποιεί μηνύματα JSON όπως έχουμε αναφέρει. Για να χρησιμοποιηθεί το Google Maps API από μία εφαρμογή απαιτείται η απόκτηση ενός "κλειδιού" (Google Maps API Key) από τον δημιουργό της εφαρμογής που το χρησιμοποιεί και η εισαγωγή του στον κώδικα της εφαρμογής.

Οι δυνατότητες του Google Maps JavaScript API που χρησιμοποιούμε στο SocialMap script είναι οι εξής:

- Markers (pinpoints).
- InfoWindow με ενεργοποίηση στο συμβάν πατήματος του marker (On Click).
- Polylines.
- Μετακίνηση του κέντρου του χάρτη.

6.2 Markers

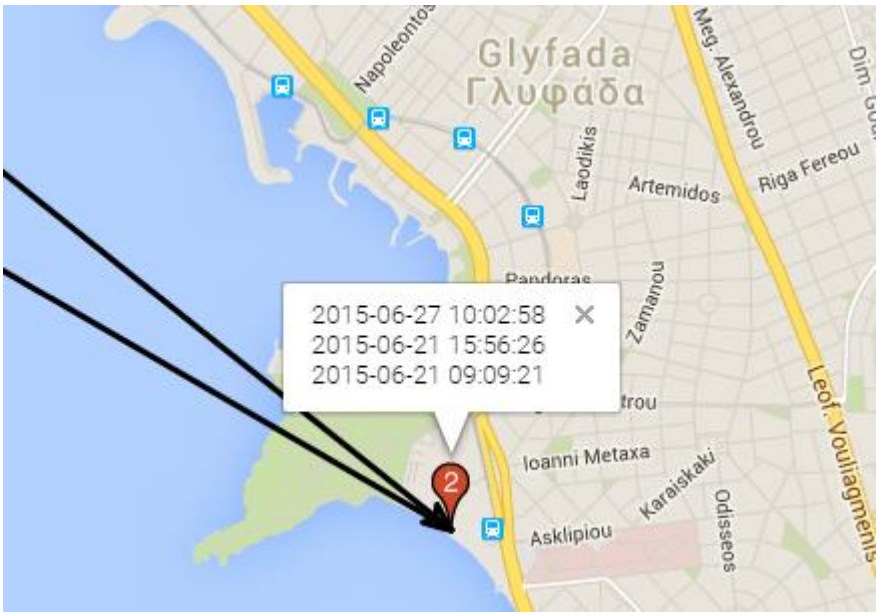
Ένα marker (pinpoint) αντιπροσωπεύει ένα σημείο ενδιαφέροντος στον χάρτη. Η εικόνα ενός marker είναι προκαθορισμένη, αλλά μπορεί να αλλάξει με διαφορετική, δικής μας σχεδίασης. Μπορούν να δεχτούν παραμέτρους όπως διαστάσεις, χρώμα, ετικέτα. Στην εφαρμογή μας καθορίζουμε το χρώμα των markers που αφορούν δραστηριότητα του Twitter με μπλε χρώμα (#094AB2) και άσπρη ετικέτα (#FFFFFF), μεγέθους 21 X 34 Pixels και την δραστηριότητα που αφορά το Instagram με κόκκινο χρώμα (#D24726) και άσπρη ετικέτα (#FFFFFF), μεγέθους 21 X 34 Pixels. Με μεγαλύτερη διάσταση 40 X 61 Pixels και πράσινο χρώμα (#008A00) και με άσπρη ετικέτα (#FFFFFF), εμφανίζεται το περισσότερο πρόσφατο σημείο δραστηριότητας και στα δύο social media.

Social Media	Χρώμα	Διαστάσεις	Εικόνα
Twitter	#094AB2	21 X 34 Pixels	
Instagram	#D24726	21 X 34 Pixels	
Πιο πρόσφατο	#008A00	40 X 61 Pixels	

Πίνακας 1 *- Απεικόνιση SocialMap markers.*

6.3 Info Window

Τα info windows [Google, 2014] χρησιμοποιούνται για να απεικονίσουν περιεχόμενο σε μορφή κειμένου σε ένα αναδυόμενο (popup) παράθυρο, πάνω σε ένα marker στον χάρτη. Ενεργοποιούνται με κάποιο συμβάν το οποίο συνήθως είναι η επιλογή με το ποντίκι (event on mouse click). Στην εφαρμογή μας κάνουμε συνάθροιση των γεωγραφικών σημείων, όπου έχουμε πολλαπλές επισκέψεις – εγγραφές, με σκοπό να βελτιστοποιήσουμε την απεικόνιση των διαδρομών. Λαμβάνουμε υπόψη μας την εγγραφή με την νεότερη ημερομηνία και απεικονίζουμε τις εγγραφές από την νεότερη στην παλαιότερη, όπως φαίνεται στην παρακάτω εικόνα.



Εικόνα 30 *- Παράδειγμα Info Window σε σημείο ενδιαφέροντος με πολλαπλές επισκέψεις - εγγραφές.*

6.4 Polylines

Το Google Maps API έχει δύο κλάσεις που διαχειρίζονται γεωμετρικά σχήματα. Είναι οι κλάσεις polylines [Google, 2014] και polygons. Οι κλάσεις αυτές μας παρέχουν τα απαραίτητα εργαλεία για να σχεδιάσουμε δρόμους, σύνορα και κατευθύνσεις. Τα polylines χρησιμοποιούνται στην εφαρμογή μας για να σχεδιάσουμε τα μονοπάτια, ανάμεσα στα σημεία δραστηριότητας του χρήστη των δύο social media, με χρονολογική σειρά.

Τα polylines ουσιαστικά είναι μια αλληλουχία συνδεδεμένων κατευθυνόμενων τόξων, ανάμεσα σε σημεία ενδιαφέροντος.

Εικόνα 31 *- Παράδειγμα χρήσης polylines.*

Στην εφαρμογή μας χρησιμοποιήσαμε επιπλέον ένα vector icon στο polyline, το οποίο ονομάζεται (FORWARD_CLOSED_ARROW) symbol [Google, 2014], έτσι ώστε να δώσουμε την κατεύθυνση στο polyline. Τα προκαθορισμένα σύμβολα του API φαίνονται στο παρακάτω σχήμα.

Name	Description	Example
google.maps.SymbolPath.CIRCLE	A circle.	
google.maps.SymbolPath. BACKWARD_CLOSED_ARROW	A backward-pointing arrow that is closed on all sides.	
google.maps.SymbolPath. FORWARD_CLOSED_ARROW	A forward-pointing arrow that is closed on all sides.	
google.maps.SymbolPath. BACKWARD_OPEN_ARROW	A backward-pointing arrow that is open on one side.	
google.maps.SymbolPath. FORWARD_OPEN_ARROW	A forward-pointing arrow that is open on one side.	

Εικόνα 32 *- Προκαθορισμένα σύμβολα στο Google API.*

6.5 Map.panTo

Στο SocialMap python script και συγκεκριμένα στον JavaScript κώδικα που περιλαμβάνεται στο HTML αρχείο, γίνεται χρήση της παραμέτρου panTo για τον Google Maps χάρτη μας. Η λογική πίσω από την συγκεκριμένη λειτουργία, είναι ότι ζητάμε από το χάρτη να κεντράρει το παράθυρο του web browser μας με βάση τις γεωγραφικές συντεταγμένες του εκάστοτε marker. Με αυτό τον τρόπο προσπαθούμε να δώσουμε την αίσθηση της κίνησης και της παρακολούθησης της κατεύθυνσης της διαδρομής στα σημεία ενδιαφέροντος.

6.6 Google maps geocoding API

Το Google maps geocoding API [Google, 2014], αναφέρεται στην λειτουργία της μετατροπής ταχυδρομικών διευθύνσεων σε γεωγραφικές συντεταγμένες, τις οποίες στην συνέχεια μπορούμε να τις απεικονίσουμε με την βοήθεια του Google maps JavaScript API, που αναλύσαμε προηγουμένως.

Σε περιπτώσεις όπου οι χρήστες των Twitter και Instagram, κάνουν χρήση της υπηρεσίας μέσω του smartphone τους, έχοντας όμως απενεργοποιήσει την λειτουργία του GPS, μπορούμε να χρησιμοποιήσουμε την παράμετρο των location tags (αντί των coordinates tags), για να πάρουμε μια λιγότερη ακριβής εκτίμηση της τοποθεσίας τους (ακρίβεια περίπου 1 Km). Με αυτόν τον τρόπο έχουμε περισσότερο αριθμό εγγραφών σχετικές με τοποθεσία, με συνέπεια να έχουμε περισσότερα σημεία ενδιαφέροντος στον χάρτη του χρήστη.

Και στην συγκεκριμένη περίπτωση η χρήση του API, συνεπάγεται περιορισμούς που αφορούν την συχνότητα και την ποσότητα των ερωτήσεων και απαντήσεων και απεικονίζονται στην παρακάτω εικόνα.

Standard Usage Limits	
Users of the standard API: • 2,500 free requests per day • 10 requests per second	Enable pay-as-you-go billing to unlock higher quotas: $0.50 USD / 1000 additional requests, up to 100,000 daily.

Εικόνα 33 *- Περιορισμοί χρήσης Google maps geocoding API.*

Κεφάλαιο 7
Python Script SocialMap

7.1 Εισαγωγή στην Python

Η γλώσσα προγραμματισμού Python είναι μια υψηλού επιπέδου γλώσσα προγραμματισμού, η οποία δημιουργήθηκε το 1990 και αναπτύσσεται ως ανοιχτό λογισμικό (open source) ενώ η διαχείρισή της γίνεται από τον μη κερδοσκοπικό οργανισμό Python Software Foundation. Ο κώδικας διανέμεται με την άδεια Python Software Foundation License η οποία είναι συμβατή με την GPL. Το όνομα της γλώσσας προέρχεται από την ομάδα Άγγλων κωμικών Μόντυ Πάιθον [Python, 2015]. Ο κύριος στόχος της είναι η αναγνωσιμότητα του κώδικά της και η ευκολία χρήσης της, ενώ το συντακτικό της επιτρέπει στους προγραμματιστές να εκφράσουν έννοιες σε λιγότερες γραμμές κώδικα απ' ότι θα ήταν δυνατόν σε άλλες δημοφιλής γλώσσες όπως η C++ ή η Java. Είναι ιδιαίτερα διαδεδομένη στους τομείς του penetration testing και hacking λόγω των παρακάτω χαρακτηριστικών της.

- **Απλή**: Η ομοιότητα της Python με ψευδοκώδικα είναι ένα από τα πιο ισχυρά σημεία της. Είναι μια απλή και μινιμαλιστική γλώσσα.
- **Ελεύθερη και Ανοικτού Κώδικα Γλώσσα υψηλού επιπέδου**: Η Python είναι ένα παράδειγμα Ελεύθερου Λογισμικού και Λογισμικό Ανοικτού Κώδικα, με μεγάλη ενεργή κοινότητα και συμβατότητα σε Windows / Linux / BSD / Mac OS λειτουργικά.
- **Εκτεταμένο σετ βιβλιοθηκών**: Η βιβλιοθήκη της Python είναι πραγματικά τεράστια.

Η γλώσσα προγραμματισμού Python είναι η προτιμώμενη επιλογή των κυβερνοεγκληματιών, εξαιτίας της απλότητας της, της αποτελεσματικότητας της και της

μεγάλης ποικιλίας των εξωτερικών βιβλιοθηκών της. Επιπλέον είναι διαθέσιμη σε όλες τις πλατφόρμες (Windows, Linux, MacOS), αποτελώντας εξαιρετικό προγραμματιστικό εργαλείο. Υπάρχει διαθέσιμος μεγάλος όγκος βιβλιογραφίας και ταυτόχρονα μεγάλος κύκλος ενεργών προγραμματιστών, σε σχέση με άλλες αντίστοιχες γλώσσες όπως η Ruby. Η επεκτασιμότητα της Python έναντι άλλων γλωσσών είναι δεδομένη. Η Python επιπλέον είναι πιο γρήγορη στην εκτέλεση του κώδικα από ότι για παράδειγμα η Ruby. Στον παρακάτω πίνακα γίνεται μια σύγκριση ανάμεσα σε Python και Ruby.

	Python	**Ruby**
Χαρακτηριστικό γλώσσας	περισσότερο άμεση και εύκολη	δίνει απόλυτη ελευθερία και δύναμη στους προγραμματιστές
Πλεονεκτήματα	πολύ έυκολη εκμάθηση, μεγάλη κοινότητα συνδεδεμένη με το Linux	πολλές δυνατότητες όσον αφορά το web development, μεγαλύτερη ευελιξία
Μειονεκτήματα	μεγαλύτερη δυσκολία στην δημιουργία web apps	δύσκολη αποσφαλμάτωση
Web frameworks	Django	Ruby on Rails
Κοινότητα	Σταθερή και ευρεία	Προσφέρει περισσότερη καινοτομία, αλλά όχι τόσο σταθερή και με επίκεντρο το web dev
Από ποιους χρησιμοποιείται	Google, Pinterest, Instagram, Mozilla Firefox	Hulu, Groupon, Twitter, GitHub, Shopify

Πίνακας 2 *- Σύγκριση Python και Ruby*

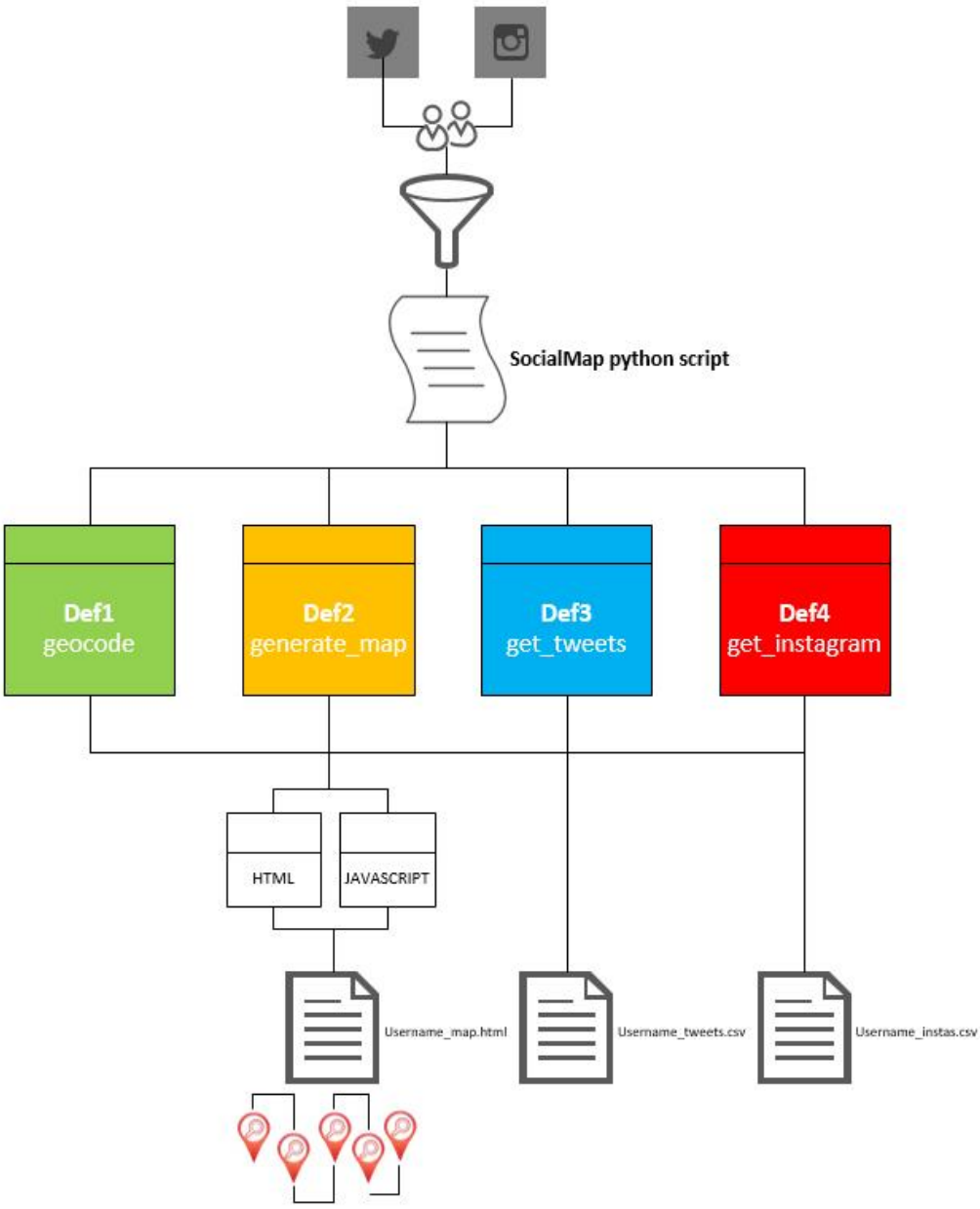
7.2 SocialMap python script βασικές λειτουργίες

Οι βασικές λειτουργίες που μας ενδιέφερε να πραγματοποιεί το python πρόγραμμά μας είναι:

- Ανάγνωση 1000 τελευταίων tweets από τον λογαριασμό twitter του χρήστη-στόχου. Σύνδεση μέσω twitter API και αποθήκευση σε tab delimited csv Unicode αρχείου, του timestamp, twit κειμένου και του πεδίου geolocation.

- Ανάγνωση 200 τελευταίων instas από τον λογαριασμό Instagram του χρήστη-στόχου. Σύνδεση μέσω Instagram API και αποθήκευση σε tab delimited csv Unicode αρχείου, του timestamp, insta κειμένου και του πεδίου geolocation.
- Απεικόνιση των twits που περιέχουν μετά-δεδομένα geolocation (GPS long, lat) ή και location tag (όχι ακριβές στίγμα, αλλά περιοχή π.χ. New York, USA) σε Google maps χάρτη. Τα γεωγραφικά δεδομένα απεικονίζονται από το παλαιότερο στο πιο πρόσφατο με αριθμημένα pinpoints με μπλε χρώμα (το πιο πρόσφατο είναι το νο.1 με πράσινο χρώμα).
- Απεικόνιση των instas που περιέχουν μετά-δεδομένα geolocation (GPS long, lat) ή και location tag (όχι ακριβές στίγμα, αλλά περιοχή π.χ. New York, USA) σε Google maps χάρτη. Τα γεωγραφικά δεδομένα απεικονίζονται από το παλαιότερο στο πιο πρόσφατο με αριθμημένα pinpoints με κόκκινο χρώμα (το πιο πρόσφατο είναι το νο.1 με πράσινο χρώμα).
- Απεικόνιση των εγγραφών που αφορούν και τα δύο social networks σε Google maps. Δημιουργία δυνατότητας στον χάρτη travel by map. Τα σημεία δραστηριότητας ενώνονται μεταξύ τους με διανυσματικά βέλη που σχεδιάζονται πάνω στον χάρτη και «ταξιδεύουν» από το παλαιότερο σημείο παρουσίας στο πιο πρόσφατο.
- Clustered pinpoints για γεωγραφικά σημεία που έχουμε πολλαπλές εγγραφές. Σε information box φαίνονται όλα τα timestamps των εγγραφών. Για παράδειγμα εάν στην τοποθεσία pinpoint 1 έχουν συγκεντρωθεί 40 tweets τότε το επόμενο pinpoint θα έχει τον αριθμό 41 και όχι 2. Οι εγγραφές για τα pinpoints είναι συγκεντρωτικές όχι όμως και οι διαδρομές (paths) που φαίνονται μια προς μια.

Στην παρακάτω εικόνα αναφέρονται διαγραμματικά οι λειτουργίες του SocialMap python script.



Εικόνα 34 *- Block διάγραμμα του SocialMap python script.*

Εάν θέλουμε να δώσουμε μια σύντομη περιγραφή του στόχου του συγκεκριμένου python προγράμματος είναι να αποδείξουμε την ερευνητική υπόθεση, ότι κάποιος κακόβουλος χρήστης, αξιοποιώντας τα γεωγραφικά δεδομένα που μπορεί να ανακαλέσει από τις social media πλατφόρμες twitter και Instagram για έναν στόχο χρήστη, μπορεί να ξέρει σχεδόν σε πραγματικό χρόνο την τοποθεσία και διαδρομή του στόχου σε πάνω σε χάρτη. Και μάλιστα ανώνυμα και χωρίς να αφήσει ίχνη.

Για την υλοποίηση των παραπάνω δυνατότητων, λειτουργιών στο python πρόγραμμα θα χρησιμοποιηθούν:

- Twitter API
- Instagram API
- Google geocode API
- Google maps API
- Tweepy python βιβλιοθήκη
- Instagram python βιβλιοθήκη
- Csv python βιβλιοθήκη
- Json python βιβλιοθήκη
- Codecs python βιβλιοθήκη
- Geopy python βιβλιοθήκη
- Html file για την παρουσίαση του χάρτη
- Javascript ενσωματωμένη στο html για την εμφάνιση των pinpoints και των polyline (travel by map)

Τα διαφορετικά APIs έχουν αναφερθεί σε προηγούμενα κεφάλαια, οπότε θα επικεντρωθούμε στην λογική των επιμέρους βιβλιοθηκών και τεχνικών για την επίτευξη του στόχου μας.

7.3 Tweepy python βιβλιοθήκη

Το tweepy python library [Tweepy, 2015] είναι μια εύκολη στην χρήση βιβλιοθήκη που εξασφαλίζει την διασύνδεση με το Twitter API μέσω της python. Το API class μας παρέχει πρόσβαση στην πλήρη λίστα μεθόδων του twitter RESTful API. Κάθε μέθοδος μπορεί να δεχτεί παραμέτρους και να επιστρέψει αποτελέσματα. Στο πρόγραμμά μας όταν επικαλούμαστε ένα API method, αυτό που επιστρέφεται είναι ένα model class instance, που θα περιλαμβάνει τα δεδομένα από το twitter.

Η πιστοποίηση που υποστηρίζει το tweepy είναι μέσω OAuth authentication. Το OAuth είναι ένα πρωτόκολλο ανοιχτού κώδικα το οποίο χρησιμοποιείται στην πιστοποίηση χρηστών και προγραμμάτων. Πρακτικά εξασφαλίζει εξουσιοδότηση πρόσβασης σε

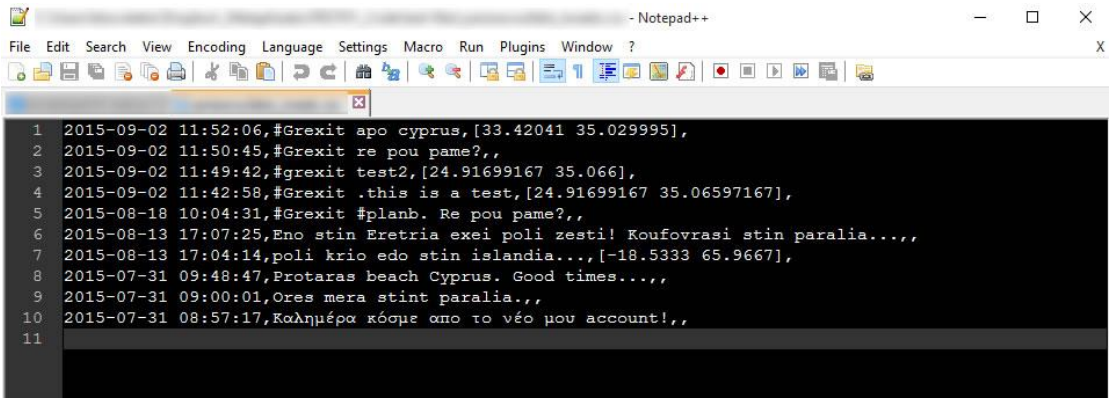
πόρους ενός συστήματος, για λογαριασμό του εκάστοτε ιδιοκτήτη / χρήστη. Με αυτό τον τρόπο γίνεται χρήση πόρων για λογαριασμό του χρήστη, χωρίς να γίνεται διαμοιρασμός των διαπιστευτηρίων (username / password) του εκάστοτε χρήστη. Ουσιαστικά το OAuth επιτρέπει την δημιουργία tokens που θα εκδοθούν στους πελάτες τρίτων από έναν διακομιστή άδειας, με την έγκριση του ιδιοκτήτη των πόρων.

Πρόκειται για εξωτερική βιβλιοθήκη και για να την χρησιμοποιήσουμε θα πρέπει πρώτα να την εγκαταστήσουμε στο σύστημά μας μέσω του Python package manager **pip** (pip install tweepy).

7.4 CSV python βιβλιοθήκη

Στην περίπτωση αυτή έχουμε να κάνουμε με μια βιβλιοθήκη της Python [Python, 2015] που ήδη περιλαμβάνεται στο οικοσύστημα μας και έτσι δεν χρειάζεται να γίνει εγκατάσταση, για να χρησιμοποιηθεί. Μας δίνει την δυνατότητα να γράψουμε και να διαβάσουμε μέσα από το πρόγραμμά μας αρχεία τύπου csv, έτσι ώστε να μπορέσουμε να αποθηκεύσουμε αποτελέσματα από τις αναζητήσεις από τα επιμέρους API. Στο SocialMap script έχει γίνει πρόβλεψη έτσι ώστε το csv να αποθηκεύεται σε UTF-8 τύπο. Με αυτόν τον τρόπο αποτυπώνονται σωστά οι χαρακτήρες του ελληνικού αλφαβήτου.

Το CSV αρχείο πρακτικά είναι ένα αρχείο κειμένου που ακολουθεί μια συγκεκριμένη γραμμογράφηση στην δομή του. Κάθε γραμμή αποτελεί μια νέα εγγραφή, ενώ κάθε πεδίο διαχωρίζεται με έναν ειδικό χαρακτήρα που στην δική μας περίπτωση είναι το κόμμα (,). Στην παρακάτω εικόνα φαίνεται ένα απόσπασμα από δεδομένα που έχουν συλλεγεί από το twitter ενός εικονικού χρήστη. Τα πεδία που συλλέγονται είναι: χρονοσφραγίδα, κείμενο tweet, GPS συντεταγμένες.



Εικόνα 35 *- Δοκιμαστικό αρχείο csv που παράγει το SocialMap script για το twitter.*

7.5 JSON python βιβλιοθήκη

Το JSON (JavaScript Object Notation) είναι ένα ελαφρύ πρότυπο ανταλλαγής δεδομένων. Είναι εύκολο για τον χρήστη να το διαβάσει και να το γράψει. Είναι εύκολο για τις μηχανές να το αναλύσουν (parse) και να το παράγουν (generate). Είναι βασισμένο πάνω σε ένα υποσύνολο της γλώσσας προγραμματισμού JavaScript. Το JSON είναι χτισμένο σε δύο δομές. Αφ' ενός μια συλλογή από ζευγάρια ονομάτων/τιμών. Σε διάφορες γλώσσες προγραμματισμού, αυτό αντιλαμβάνεται ως ένα object, καταχώριση, δομή, λεξικό, πίνακα hash (hash table), λίστα κλειδιών, ή associative πίνακα. Αφετέρου μία ταξινομημένη λίστα τιμών. Στις περισσότερες γλώσσες προγραμματισμού, αυτό αντιλαμβάνεται ως ένας πίνακας (array), διάνυσμα, λίστα, ή ακολουθία. Έχει περιγραφεί πλήρως με το RFC 7159 της Internet Engineering Task Force (IETF) [Force, 2014, Json, 2014]. Χρησιμοποιώντας την συγκεκριμένη ενσωματωμένη βιβλιοθήκη στο python script μας, απομονώνουμε τα πεδία που μας ενδιαφέρουν από τα social media APIs και αποθηκεύουμε τα αποτελέσματα στην συνέχεια σε csv αρχεία.

Εικόνα 36 *- JSON data file για το Twitter [Twitter IDs, 2014]*

7.6 Codecs python βιβλιοθήκη

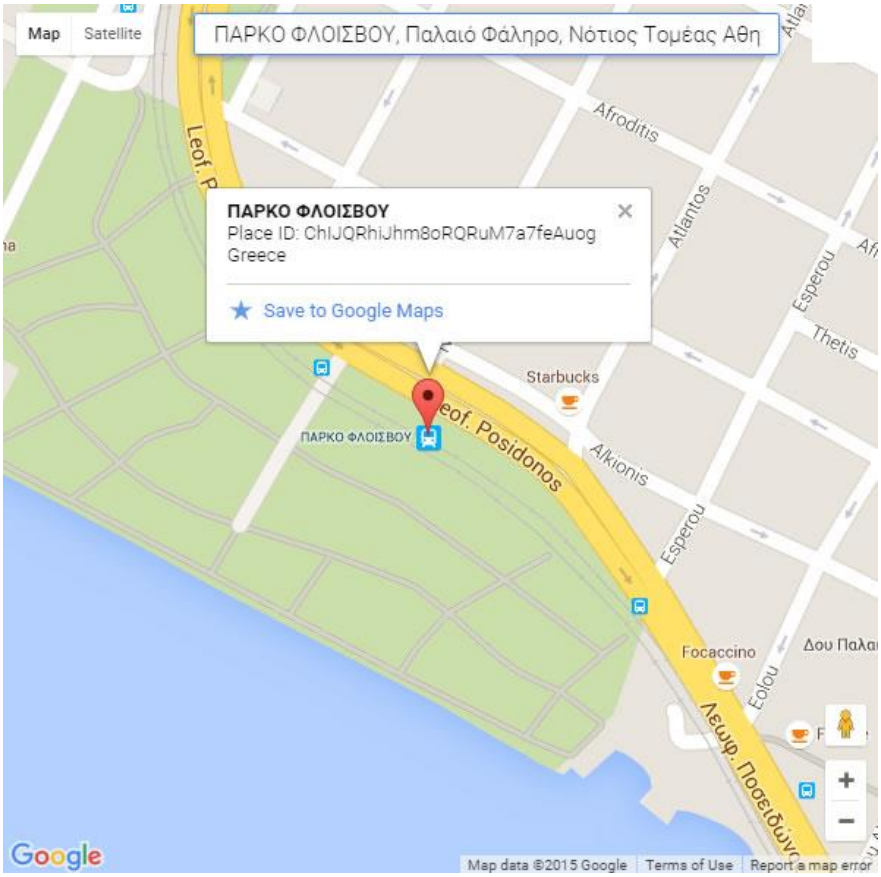
Η ενσωματωμένη python βιβλιοθήκη codecs [Python, 2014], μας παρέχει την δυνατότητα να διαβάσουμε ή να γράψουμε ένα δομημένο αρχείο, που στην περίπτωση μας είναι το csv αρχείο. Χρησιμοποιείται για την σωστή ανάγνωση και εγγραφή UTF-8 αρχείων.

7.7 Geopy python βιβλιοθήκη

Η geopy [Geopy, 2014] είναι μια βιβλιοθήκη για την Python, που προσφέρει διασύνδεση με διάφορες υπηρεσίες geocoding, όπως οι OpenStreetMap Nominatim, ESRI ArcGIS, Google Geocoding API (V3), Baidu Maps, Bing Maps API, Yahoo! PlaceFinder, Yandex, IGN France, GeoNames, NaviData, OpenMapQuest, What3Words, OpenCage, SmartyStreets, geocoder.us, GeocodeFarm. Πρόκειται για εξωτερική βιβλιοθήκη και για να την χρησιμοποιήσουμε θα πρέπει πρώτα να την εγκαταστήσουμε στο σύστημά μας μέσω του Python package manager pip (pip install geopy). Στην περίπτωση του SocialMap python script χρησιμοποιούμε την υπηρεσία Google Geocoding API (V3) [Geocoding, 2014].

Το geocoding αναφέρεται ως η διαδικασία της μετατροπής μιας διεύθυνσης (π.χ. Λεωφ. Γιάννου Κρανιδιώτη 33, 2220, Λατσιά, Κύπρος), σε γεωγραφικές συντεταγμένες (στο παράδειγμά μας (lat, long) = (35.114578, 33.377240)), τις οποίες τελικά μπορείς να τις αναπαραστήσεις σε σημεία πάνω σε χάρτη. Στο script μας όπου ο χρήστης μας δεν αποκαλύπτει τις πραγματικές συντεταγμένες από το smartphone του μέσω του tag **coordinates**, τότε σαν εναλλακτικό τρόπο προσδιορισμού της τοποθεσίας του, χρησιμοποιούμε το tag **location** μέσω του οποίου δηλώνουμε την τοποθεσία, πόλη που βρισκόμαστε.

Αντίστροφα το reverse geocoding είναι η διαδικασία μετατροπής γεωγραφικών συντεταγμένων ή place ID σε μια ταχυδρομική διεύθυνση. Η πρόσβαση στο Google geocoding γίνεται μέσω HTTP request και η απάντηση είναι για άλλη μια φορά σε μορφή JSON.



Εικόνα 37 *- Google Place ID για το Πάρκο Φλοίσβου, Παλαιό Φάληρο, Αθήνα, Ελλάδα.*

7.8 HTML αρχείο

Το SocialMap python script αποθηκεύει σαν αρχείο απεικόνισης του χάρτη και του εικονικού ταξιδιού μέσω του χάρτη, ένα αρχείο html όπου με χρήση JavaScript κώδικα απεικονίζονται σε pin points τα σημεία παρουσίας του χρήστη των social media, τα οποία έχουμε αποκαλύψει και αποθηκεύσει σε csv εκ των προτέρων. Ο κώδικας HTML χωρίς τον αντίστοιχο JavaScript κώδικα φαίνεται στην παρακάτω εικόνα.

```
<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Transitional//EN" "http://www.w3.org/TR/xhtml1/DTD/xhtml1-transitional.dtd">
<html xmlns="http://www.w3.org/1999/xhtml">
    <head>
        <title>SocialMap Python Script</title>
        <meta name="viewport" content="width=device-width, user-scalable=no">
        <style>
            html,body
            {
                height:100%;
                padding:0px;
                margin:0px;
            }
            #map-canvas
            {
                height:100%;
            }
        </style>
        <script type="text/javascript">
            ....
        </script>
    </head>
    <body>
        <div id="map-canvas">
        </div>
    </body>
</html>
```

Εικόνα 38 *- Το HTML αρχείο του SocialMap.*

Το αρχείο HTML είναι πολύ απλό, καθότι περιλαμβάνει τα απαραίτητα tags του HTML, ένα απλό CSS styling (γραμμές 6 – 17), ενώ την πραγματική δουλειά αναλαμβάνουν τα επιμέρους JavaScripts. Έχει γίνει ορισμός ενός div element με όνομα map-canvas όπου θα εμφανίζεται ο χάρτης μας. Τα διπλά % στην παραπάνω εικόνα αφορούν τον χαρακτήρα διαφυγής στην python.

7.9 JavaScript κώδικας ενσωματωμένος στο αρχείο HTML

Μέσα στο αρχείο HTML, περιλαμβάνονται τρεις συναρτήσεις JavaScript, που είναι υπεύθυνες για την εμφάνιση του Google χάρτη, την δημιουργία των σημείων πάνω στο χάρτη που αντιπροσωπεύουν την αντίστοιχη δραστηριότητα στα δυο social media καθώς και την δημιουργία του εφέ της μετακίνησης από το παλαιότερο χρονικά σημείο στο πιο πρόσφατο. Η ενσωμάτωση CSS styling καθώς και των JavaScripts scripts μέσα στο ίδιο HTML αρχείο, έγινε για λόγους απλότητας και μείωσης του αριθμού των συνολικών διαχειρισίμων αρχείων από το python script.

Ξεκινώντας από την απλούστερη συνάρτηση από το τέλος του JavaScript κώδικα (γραμμές 191 – 202 από το κώδικα του SocialMap python script που βρίσκεται στο

παράρτημα Κώδικας με σχολιασμό, στο τέλος της παρούσας μεταπτυχιακής διατριβής), συναντάμε την **sortArray(ary)**.

```
//function to sort the merged list with respective to the date-time
                    function sortArray(ary){
                        ary.sort(function(x,y){
                            var xd = Date.parse(x[0]);
                            var yd = Date.parse(y[0]);
                            if(xd < yd) return 1;
                            if(xd > yd) return -1;
                            return 0;
                        });
                        return ary;
                    }
```

Η συνάρτηση sortArray αναλαμβάνει να ταξινομήσει την λίστα με τις εγγραφές με βάση την χρονοσφραγίδα από την παλαιότερη στην πιο πρόσφατη.

Αμέσως μετά ακολουθεί η εντολή που αφορά το HTML DOM (Document Object Model) [Hallaraker, 2005], για να ενεργοποιήσει την κεντρική JavaScript συνάρτηση **initialize()** κατά την ενεργοποίηση (load) της σελίδας HTML.

```
google.maps.event.addDomListener(window, 'load', initialize);
```

Η συνάρτηση **draw(from,to)**, αναλαμβάνει να σχεδιάσει με την μορφή ενός τόξου την διαδρομή ανάμεσα σε δύο σημεία, με κατεύθυνση από το παλαιότερο στο αμέσως νεότερο χρονικά. Για την επίτευξη του αποτελέσματος, χρησιμοποιούμε το Google Maps JavaScript API [Google, 2014] και συγκεκριμένα την λειτουργία σχεδιασμού πάνω στο χάρτη με Polylines (forward closed arrow) [Google, 2014]. Το κατευθυντικό τόξο σχεδιάζεται στον χάρτη με βάση τις συντεταγμένες ανάμεσα στα δύο χρονικά ταξινομημένα σημεία. Ένα τέτοιο παράδειγμα απεικόνισης εμφανίζεται στην Εικόνα 32.

```
//function to draw the arrow
function draw(from,to)
{
 var lineSymbol =
 {
  path: google.maps.SymbolPath.FORWARD_CLOSED_ARROW
 };

// Create the polyline and add the symbol via the 'icons' property.
 var line = new google.maps.Polyline
 ({
```

```
  path: [{lat: from.lat, lng: from.lng}, {lat: to.lat, lng: to.lng}],
  icons: [{icon: lineSymbol, offset: '100%%'}],
  map: map
 });
 map.panTo({lat: to.lat, lng: to.lng});
}
```

Η τελευταία γραμμή περιλαμβάνει την εντολή panTo που αναλαμβάνει να αλλάξει το κέντρο του χάρτη, προσδίδοντας έτσι την αίσθηση της μετακίνησης του χάρτη ανάμεσα στα σημεία, μέσω του σχεδιασμού του τόξου.

Την ευθύνη για την δημιουργία, αποθήκευση και επεξεργασία των σημείων πάνω στον χάρτη έχει η μεγαλύτερη σε μέγεθος JavaScript συνάρτηση **initialize()**. Περιλαμβάνει τις παρακάτω επιμέρους διαδικασίες:

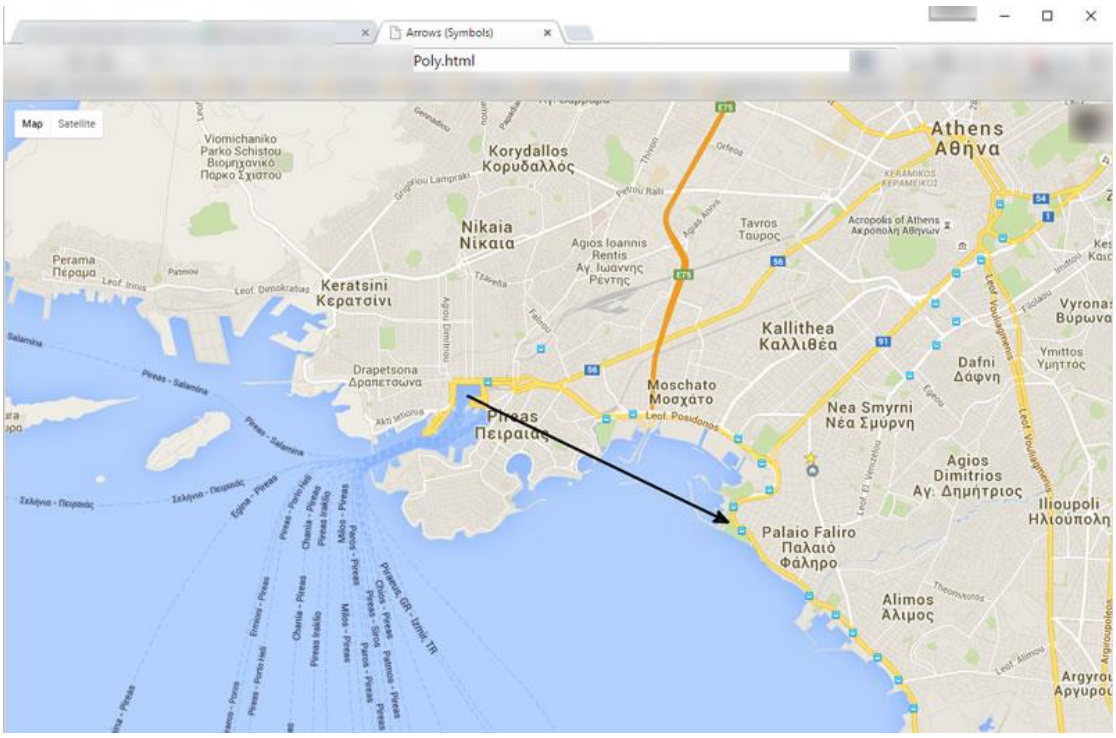
- Αποθήκευση των δεδομένων των σημείων στο HTML αρχείο, μέσω της μεταβλητής τύπου array **myLatLng**. Περιέχει πεδία για την χρονοσφραγίδα, γεωγραφικές συντεταγμένες και το γραφικό pinpoint σε URL που καλείται μέσω Google API. Επιπλέον το pinpoint έχει δεχτεί ιδιότητες χρώματος (μπλε για twitter, κόκκινο για Instagram και πράσινο για το νεότερο χρονικά) και αριθμού αλληλουχίας (sequence number). Ο μέγιστος αριθμός των εγγραφών που μπορούν να αποθηκευτούν στην συγκεκριμένη μεταβλητή έχει οριστεί σε 100. Η μορφή πραγματικών δεδομένων φαίνεται στην Εικόνα 34.
- Δημιουργία pinpoints. Διαβάζουμε όλες τις εγγραφές και μόλις συναντήσουμε γεωγραφικές συντεταγμένες που δεν έχουμε συναντήσει, εισάγουμε ένα νέο pinpoint. Εάν οι συγκεκριμένες συντεταγμένες έχουν εμφανιστεί προηγουμένως, τότε δεν δημιουργείται νέο pin point, αλλά απλά φτιάχνουμε μια νέα εγγραφή χρονοσφραγίδας στο balloon info παράθυρο (Εικόνα 33).
- Έλεγχος για το τελευταίο, με την νεότερη χρονοσφραγίδα pinpoint, το οποίο είναι μεγαλύτερο σε διαστάσεις και πράσινου χρώματος για να ξεχωρίζει.
- Δημιουργία του παραθύρου με τις πληροφορίες για τις χρονοσφραγίδες του σημείου. Ενεργοποίηση του παραθύρου με πάτημα πάνω στο pinpoint.
- Ορισμός χρονικής καθυστέρησης 0,5 δευτερολέπτου, για τον σχεδιασμό του επόμενου pinpoint.

Η javascript συνάρτηση initialize() περιλαμβάνει τον παρακάτω κώδικα:

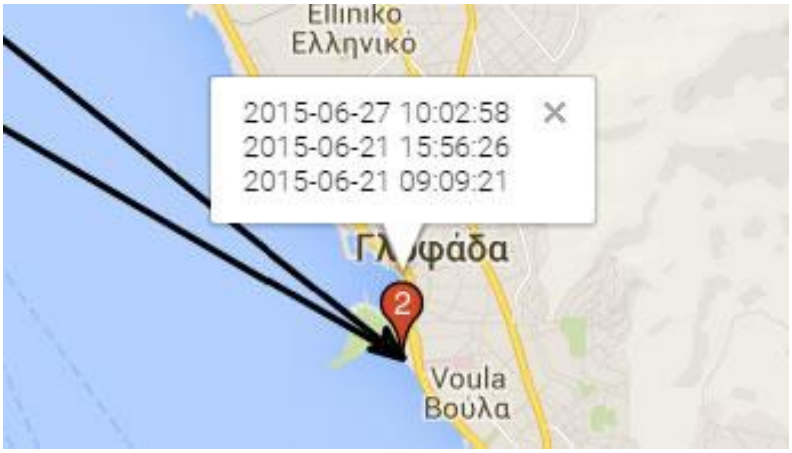
```
function initialize() {
```

```
                                //stores the marker data
                                var myLatLng = [%s];
                                //initialize the map
                                  window.map = new
google.maps.Map(document.getElementById('map-canvas'), {
                                    zoom: 8,
                                    center: myLatLng[0][1]
                                  });

                                var LIMIT = 100;
                                var maxlen = (myLatLng.length > 100) ? 100 :
myLatLng.length;

                                //sort the array by date
                                myLatLng = sortArray(myLatLng);

                                //this holds all the markers we create
                                var markerset = [];

                                //loop all the marker data
                                for(var i=0;i<maxlen;i++){
                                    var found = false;
                                    //first check if the marker is already
positioned in current gps location
                                    for(var j=0;j<markerset.length;j++){
                                        //if positioned we simply add the time
data to existing marker
                                        if(markerset[j].lat ==
myLatLng[i][1].lat && markerset[j].lng == myLatLng[i][1].lng){
                                            var marker = markerset[j];
                                            marker.info.content =
marker.info.content + "<br />" + myLatLng[i][0];
                                            //markerset.push(marker);
                                            found = true;
                                            break;
                                        }
                                    }

                                  //if not a new marker is created
                                  if(!found){
                                     //again a crude checking for green pin,
this has to be done since merging is done
                                     if(myLatLng[i][2].indexOf("008A00") > -
1){
                                        var pinIcon = new
google.maps.MarkerImage(
                                            myLatLng[i][2],
                                            null, /* size is determined at
runtime */
                                            null, /* origin is 0,0 */
                                            null, /* anchor is bottom center of
the scaled image */
                                            new google.maps.Size(40, 61)
                                        );
                                     }else{
                                         var pinIcon = new
google.maps.MarkerImage(
                                            myLatLng[i][2],
                                            null, /* size is determined at
runtime */
                                            null, /* origin is 0,0 */
```

```
                                        null, /* anchor is bottom center of
the scaled image */
                                        new google.maps.Size(21, 34)
                                    );
                                }

                                //add new marker
                                var marker = new google.maps.Marker({
                                  position: myLatLng[i][1],
                                  map: map,
                                  icon: pinIcon
                                });

                                //store the lat,lng to use in the check-
for-duplicates code
                                marker.lat = myLatLng[i][1].lat;
                                marker.lng = myLatLng[i][1].lng;

                                //adding the info window
                                marker.info = new
google.maps.InfoWindow({
                                  content: myLatLng[i][0]
                                });

                                //adding the listener to show the window
on click
                                google.maps.event.addListener(marker,
'click', function() {
                                    this.info.open(map, this);
                                 });

                                 //add the marker to the set
                                 markerset.push(marker);
                            }
                        }

                        //drawing all the markers with .5s interval
                        var pointer = markerset.length - 1;
                        window.drawing = setInterval(function(){
                            draw(markerset[pointer],markerset[pointer-
1]);

                            pointer--;
                            if(pointer == 0){
                                clearInterval(drawing);
                            }
                        },500);


                    }
```



Εικόνα 39 *- Google Maps Javascript polyline arrow παράδειγμα.*



Εικόνα 40 *- Συνάθροιση εγγραφών σε Google pinpoint Info balloon*

```
//stores the marker data
var myLatLng = [["2015-07-15 16:21:04",{ lat: 39.5899629, lng: 19.899001 }, "
http://chart.apis.google.com/chart?chst=d_map_pin_letter&chld=1|008A00|FFFFFF"],[
"2015-06-27 10:02:58",{ lat: 37.849556311, lng: 23.75062235 }, "
http://chart.apis.google.com/chart?chst=d_map_pin_letter&chld=2|D24726|FFFFFF"],[
"2015-06-23 15:47:49",{ lat: 39.665477024, lng: 20.840999688 }, "
http://chart.apis.google.com/chart?chst=d_map_pin_letter&chld=3|D24726|FFFFFF"],[
"2015-06-21 15:56:26",{ lat: 37.849556311, lng: 23.75062235 }, "
http://chart.apis.google.com/chart?chst=d_map_pin_letter&chld=4|D24726|FFFFFF"],[
"2015-06-21 09:09:21",{ lat: 37.849556311, lng: 23.75062235 }, "
http://chart.apis.google.com/chart?chst=d_map_pin_letter&chld=5|D24726|FFFFFF"],[
"2015-04-26 14:25:29",{ lat: 36.4333, lng: 28.2167 }, "
http://chart.apis.google.com/chart?chst=d_map_pin_letter&chld=6|D24726|FFFFFF"],[
"2015-04-12 19:09:00",{ lat: 51.437554018, lng: 7.795546493 }, "
http://chart.apis.google.com/chart?chst=d_map_pin_letter&chld=7|D24726|FFFFFF"],[
"2015-02-01 06:00:25",{ lat: 37.976057613, lng: 23.750193037 }, "
http://chart.apis.google.com/chart?chst=d_map_pin_letter&chld=8|D24726|FFFFFF"],[
"2015-01-02 23:36:03",{ lat: 37.977448832, lng: 23.741272351 }, "
http://chart.apis.google.com/chart?chst=d_map_pin_letter&chld=9|D24726|FFFFFF"],[
"2015-01-01 01:01:48",{ lat: 37.985780957, lng: 23.720178008 }, "
http://chart.apis.google.com/chart?chst=d_map_pin_letter&chld=10|D24726|FFFFFF"],];
```

Εικόνα 41 *- Πραγματικά δεδομένα που αντιστοιχούν σε pinpoint δραστηριότητας στα δύο social media.*

Κεφάλαιο 8
Νομικά και ηθικά θέματα

8.1 Νομοθεσία για την προστασία δεδομένων προσωπικού χαρακτήρα

Το 1997 το ελληνικό Κοινοβούλιο θέσπισε με τον νόμο Ν.2472/97 περί ιδιωτικότητας «για την προστασία του ατόμου έναντι της επεξεργασίας δεδομένων προσωπικού χαρακτήρα». Θεμέλιο της νομοθετικής παρέκβασης είναι το δικαίωμα ανθρώπων να γνωρίζουν ποιος, τι και με ποιο σκοπό επεξεργάζεται πληροφορίες που τους αφορούν και να (συν)αποφασίζουν καταρχήν οι ίδιοι ποιες προσωπικές πληροφορίες θα γίνουν γνωστές στο περιβάλλον τους. Το σύστημα προστασίας που εισήγαγε το Κοινοβούλιο με την ψήφιση του νόμου 2472/97 βασίζεται στις ακόλουθες αρχές: η επεξεργασία προσωπικών πληροφοριών είναι επιτρεπτή μόνο στις περιπτώσεις που ο νόμος προσδιορίζει περιοριστικά και δεσμευτικά, η επεξεργασία επιτρέπεται μόνο για νόμιμους, θεμιτούς και εξειδικευμένους σκοπούς που είναι γνωστοί στον πολίτη, αναγνωρίζονται και κατοχυρώνονται νέα δικαιώματα των πολιτών για να αμύνονται έναντι των προσβολών της ιδιωτικής ζωής και της προσωπικότητάς τους (δικαίωμα προηγούμενης πληροφόρησης, διόρθωσης, αποζημίωσης), ιδρύθηκε η Αρχή Προστασίας Προσωπικών δεδομένων με αντικείμενο τον έλεγχο της τήρησης της σχετικής νομοθεσίας και ευρύτατες αρμοδιότητες [Α.Π.Δ.Π.Χ, 2009]. Οι ρυθμίσεις του Ν.2472/97 συμπληρώθηκαν από τον Ν. 2774/99 για την «Προστασία δεδομένων προσωπικού χαρακτήρα στον τηλεπικοινωνιακό τομέα». Ο νόμος αυτός κατοχύρωσε σημαντικά δικαιώματα των συνδρομητών και χρηστών τηλεπικοινωνιακών υπηρεσιών όπως το δικαίωμα να απορρίπτουν μη ζητηθείσες κλήσεις ή να μην περιλαμβάνονται σε ηλεκτρονικούς καταλόγους εφόσον δεν το επιθυμούν.

Οι εφαρμογές κοινωνικής δικτύωσης (π.χ. Facebook, Flickr, Twitter) έχουν τον ταχύτερα αναπτυσσόμενο μηχανισμό αναφορικά με την ανταλλαγή προσωπικών και

επαγγελματικών πληροφοριών. Το μεγάλο ποσοστό των νέων που κάνουν χρήση αυτών των μέσων επικοινωνίας, καθώς και από όλες τις ηλικιακές ομάδες, ακόμη και τους ηλικιωμένους, αναδεικνύουν τις online εφαρμογές κοινωνικής δικτύωσης ως ένα σημαντικό μέσο διαμοιρασμού φωτογραφιών, video, αλλά και προσωπικών απόψεων.

Από την στιγμή που σε μία αστυνομική έρευνα συμπεριλαμβάνεται αναζήτηση και επεξεργασία στα προσωπικά δεδομένα, τότε σε μια τέτοια περίπτωση ισχύει το δικαίωμα προστασίας της ιδιωτικής ζωής σύμφωνα με το άρθρο 8 της Ευρωπαϊκής Σύμβασης για την Προστασία των Δικαιωμάτων του Ανθρώπου και των Θεμελιωδών Ελευθεριών (ΕΣΔΑ) [Cohen, 1992:99]. Αυτή είναι η περίπτωση κατά την οποία η επεξεργασία των δεδομένων προσωπικού χαρακτήρα στο πλαίσιο των ερευνών (OSINT) έρχεται σε αντίθεση με το δικαίωμα των πολιτών στην ιδιωτικότητα.

Το δικαίωμα στην ιδιωτικότητα προστατεύει την ιδιωτική ζωή, όταν είναι ορατή στο κοινό, [Cohen, 1992:99] και γενικά καλύπτει την επεξεργασία των δεδομένων που αφορούν την ιδιωτική ζωή των ατόμων [Rodrigues, 2000:2]. Αυτό ωστόσο δεν σημαίνει, ότι το OSINT παρεμβαίνει την ιδιωτική ζωή. Η ΕΣΔΑ θεωρεί ότι η επεξεργασία δεδομένων αφορούν αποθήκευση των δεδομένων [Eysenbach, 2001:1104], αφήνοντας ανοιχτό το ζήτημα αν η αναζήτηση και μελέτη των δεδομένων θα είναι απλή, χωρίς αποθήκευση ή τη χρήση τους, συνιστά προσβολή. Επίσης, η έγκαιρη διαμόρφωση κανονισμού ότι «οι δημόσιες πληροφορίες μπορεί να εμπίπτουν στο πεδίο εφαρμογής της ιδιωτικής ζωής όπου συστηματικά συλλέγονται και αποθηκεύονται σε αρχεία που κατέχονται από τις εξουσίες [Eysenbach, 2008:5] ενώ δείχνει ότι μη-συστηματικά επεξεργασμένες δημόσια διαθέσιμων πληροφοριών, δεν εμπίπτουν κατ' ανάγκη στο πεδίο εφαρμογής της προστασίας της ιδιωτικής ζωής. Αυτό σημαίνει ότι όσο πιο τυχαία, γίνονται αναζητήσεις για τα προσωπικά δεδομένα ενός ατόμου, δεν χρειάζεται να παρεμβαίνει στο δικαίωμα του ατόμου στην προστασία της ιδιωτικής ζωής, αλλά πιο συστηματικές έρευνες θα καλύψουν πιθανό το όριο των παρεμβάσεων [Bradke, 2009:76]. Βεβαίως, ακόμη εάν υπάρχει αναζήτηση OSINT για μόνο όταν γίνεται αυτοματοποιημένα η συλλογή δεδομένων για ένα συγκεκριμένο άτομο αυτό θα επηρεάσει το δικαίωμα στην ιδιωτική ζωή, αν διευκολύνονται από την αυτοματοποιημένη συλλογή δεδομένων, την επιλογή και την παρουσίαση των συστημάτων OSINT τότε τα αποτελέσματα της αναζήτησης μπορούν να αποδώσουν εύκολα σημαντική εικόνα για την προσωπική ζωή του ατόμου. Όταν υπάρχει συνεχή αναζήτηση OSINT για προσωπικά δεδομένα αυτό θα αποτελέσει

αναμφισβήτητα πρόβλημα. Ως εκ τούτου, θα πρέπει να υποθέσουμε ότι, εκτός από έκτακτες και μη αυτόματες αναζητήσεις από τις αρχές, χωρίς την αποθήκευση των αποτελεσμάτων της αναζήτησης σε αρχεία, οι έρευνες OSINT ενδέχεται να επηρεάσουν το δικαίωμα στην ιδιωτικότητα.

Αυτό σημαίνει ότι η έρευνα πρέπει να πληρεί τις απαιτήσεις του άρθρου 8 παράρτημα 2 της ΕΣΔΑ. Δηλαδή η έρευνα θα πρέπει να είναι σύμφωνα με το νόμο, αναγκαία σε μια δημοκρατική κοινωνία, και προς το συμφέρον της δημόσιας ασφάλειας ή την πρόληψη των αναταραχών ή εγκληματικών πράξεων. Πιο συγκεκριμένα για τις ανακριτικές αρχές, «σύμφωνα με το νόμο» σημαίνει ότι πρέπει να υπάρξει μια νομική και θεσμική βάση. Η εξαγωγή δεδομένων ανοιχτού κώδικα γενικά δεν ρυθμίζονται από τη νομοθεσία στις περισσότερες χώρες [Hanson, 2011:174], αλλά σε ορισμένες χώρες έχουν θεσπιστεί διατάξεις σχετικά με την εξαγωγή δεδομένων για την έρευνα της εγκληματικότητας και της κατασκοπείας. Δεδομένου ότι δεν υπάρχει νομοθεσία ακόμη και για την εξαγωγή δεδομένων που πραγματοποιεί η αστυνομία είναι επίσης δύσκολο να πούμε πώς υπάρχει νομική βάση για την παράβαση της ιδιωτικότητας. Οι νόμοι πρέπει να είναι ιδιαίτερα συγκεκριμένοι σε περίπτωση που οι αστυνομικές επιχειρήσεις παραβιάζουν την ιδιωτικότητα [Parsell, 2008:45]. Ωστόσο, όσο πιο συστηματικά γίνονται τα είδη ερευνών μέσω της εξαγωγής δεδομένων, τα οποία είναι πιθανό να δώσουν μια πιο παρεμβατική εικόνα της προσωπικής ζωής των ατόμων, τόσο περισσότερο ρητή νομική βάση θα απαιτηθεί [Weitzman, 2010:4]. Χώρες στις οποίες η αστυνομία εφαρμόζει εξαγωγή δεδομένων χρησιμοποιώντας εξελιγμένο λογισμικό θα πρέπει να θεσπιστεί αντίστοιχη ειδική νομοθεσία, όπου θα καθορίζονται οι όροι και το πεδίο εφαρμογής των αστυνομικών ερευνών.

Υποθέτοντας ότι υπάρχει νομική βάση, η απαίτηση της αναγκαιότητας σε μια δημοκρατική κοινωνία, σημαίνει ότι πρέπει να εφαρμόζονται ορισμένες διασφαλίσεις, ώστε να ελαχιστοποιηθεί η παραβίαση της ιδιωτικής ζωής σε ό, τι είναι αναγκαίο για τους σκοπούς της έρευνας. Οι σχετικές διασφαλίσεις θα μπορούσε να συνεπάγεται περιορισμούς του χρόνου, επιλογή και εκπαίδευση των αστυνομικών που επιτρέπεται να διεξάγουν τέτοιου είδους έρευνες, τον περιορισμό της συγχώνευσης δεδομένων με άλλες ομάδες δεδομένων, και τους περιορισμούς στις αναζητήσεις όσο αναφορά ευαίσθητους τύπους πληροφοριών. Όσο πιο συστηματική και ολοκληρωμένη έρευνα μέσω της εξαγωγής δεδομένων επιτρέπεται να γίνετε σύμφωνα με το νόμο, τόσο ποιο ισχυρή και

διασφαλισμένη από τον νόμο θα είναι. Σε γενικές γραμμές, τα κράτη μέλη έχουν μια μάλλον ευρεία διακριτική ευχέρεια όσον αφορά τον προσδιορισμό, του τι είναι αναγκαίο σε μια δημοκρατική κοινωνία, και δεδομένου ότι έχουμε να κάνουμε με δημόσια διαθέσιμες πληροφορίες, ειδικοί νόμοι που επιτρέπουν στην αστυνομία να εξάγουν δεδομένα ανοικτού κώδικα για ορισμένους σκοπούς θα ανταποκρίνεται πιθανότατα στο άρθρου 8 παράρτημα 2 της ΕΣΔΑ.

Ένα δεύτερο θέμα είναι η δυνατότητα χρήσης ως απόδειξη των αποτελεσμάτων αναζήτησης από την εξαγωγή δεδομένων. Σε μεγάλο βαθμό, αυτό που μπορεί να χρησιμοποιηθεί ως αποδεικτικό στοιχείο του ποινικού δικαίου δεν προσδιορίζεται από την ΕΣΔΑ, αλλά από την εθνική νομοθεσία, [Association, 2001:7] η οποία θα καθορίσει τους τύπους των δεδομένων που μπορεί να χρησιμοποιηθούν ως απόδειξη, με ποιο τρόπο πρέπει να παρουσιάζονται, καθώς και ο βαθμός στον οποίο αποδεδειγμένα μπορούν να θεωρηθούν αξιόπιστα. Συχνά, τα αποτελέσματα από τις έρευνες ανοιχτού κώδικα δεν θα πρέπει να χρησιμοποιηθεί άμεσα ως αποδεικτικό στοιχείο, αλλά μάλλον ως πληροφορία που καθοδηγεί την έρευνα. Για τις χώρες που εφαρμόζουν το δόγμα «δηλητηριασμένοι καρποί δένδρου», που σημαίνει ότι εξαιρούνται αποδεικτικά στοιχεία τα οποία έχουν προκύψει από παράνομες ερευνητικές δραστηριότητες.

Το άρθρο 6 της ΕΣΔΑ, δίνει το δικαίωμα σε μια δίκαιη δίκη, ωστόσο, περιέχει ένα δόγμα σημαντικό στο οποίο αμφισβητούνται οι αποδείξεις από την εξαγωγή δεδομένων με τεχνικές OSINT εφόσον αποκτήθηκαν παράνομα. Αν δεν υπάρχει εθνική νομοθεσία που προβλέπει ρητά τη βάση για έρευνες ανοιχτού κώδικα, και μια συγκεκριμένη έρευνα θεωρείται ότι παραβιάζει την ιδιωτική ζωή, τότε το άρθρο 8 της ΕΣΔΑ μπορεί να έχει παραβιαστεί. Αυτό αναγκαστικά δεν οδηγεί σε αποκλεισμό των αποδείξεων. Όπως έχει καθορισθεί από τον Khan στο Ευρωπαϊκό Δικαστήριο Ανθρωπίνων Δικαιωμάτων (ΕΔΑΔ), ακόμη και αν το δικαίωμα της ιδιωτικής ζωής του κατηγορουμένου έχει παραβιαστεί, η δίκη εναντίον του μπορεί να είναι δίκαιη υπό το φως όλων των περιστάσεων [Gakh, 2005:123]. Σε αυτή την περίπτωση, μια συνομιλία του Khan είχε καταγραφεί κρυφά από την αστυνομία χωρίς νομική βάση. Παρά το γεγονός ότι η συζήτηση ήταν η κύρια πηγή των αποδεικτικών στοιχείων εναντίον Khan, το δικαστήριο θεώρησε τη διαδικασία στο σύνολό της δίκαιη, τονίζοντας ότι σε όλα τα στάδια της διαδικασίας, ο εναγόμενος είχε τη δυνατότητα να αμφισβητήσει τα στοιχεία και τα δικαστήρια είχαν συζητήσει αν θα αποκλείσουν ή όχι τα αποδεικτικά στοιχεία,

λαμβάνοντας υπόψη τη μη θεσμική βάση την παρακολούθηση [Davolt Sr, 1999:24]. Αυτό συνεπάγεται ότι εάν τα κράτη διεξάγουν αστυνομικές έρευνες ανοιχτού κώδικα χωρίς επαρκή νομική βάση, αυτό μπορεί να παραβιάζει το δικαίωμα στην ιδιωτική ζωή, αλλά δεν είναι απαραίτητο να οδηγήσει σε αποκλεισμό των αποδείξεων, υπό την προϋπόθεση ότι το δικαστήριο έχει τη διακριτική ευχέρεια να εξαιρεί τα στοιχεία εφόσον κρίνει ότι οδηγεί σε ουσιαστική αδικία έναντι του κατηγορουμένου.

Ένα άλλο στοιχείο, από τις προερχόμενες πληροφορίες ως αποδεικτικό στοιχείο, μέσω χρήσης OSINT, είναι κατά πόσο οι πληροφορίες μπορούν να προσβληθούν στο δικαστήριο. Σε περίπτωση που η αναζήτηση OSINT βασίζεται σε δημοσιεύσεις του διαδικτύου που έχουν σημασία για την απόδειξη μιας υπόθεσης, τα στοιχεία της έρευνας μπορούν κανονικά να χρησιμοποιηθούν μόνο ως αποδεικτικό στοιχείο, εφόσον παράγονται με την παρουσία του κατηγορουμένου σε δημόσια ακρόαση με σκοπό την διεξαγωγή της δίκης με την συμμετοχή όλων των διαδίκων [Thompson, 2011:9]. Για στοιχεία που αποστέλλονται από τον κατηγορούμενο, η ενάγουσα αρχή θα πρέπει να είναι σε θέση να αμφισβητήσει τη δήλωση αυτή. Σε περίπτωση που τα αποδεικτικά στοιχεία αποστέλλονται από τρίτους, η ενάγουσα αρχή θα πρέπει να είναι σε θέση να αμφισβητήσει και να ανακρίνουν τον πάροχο του περιεχομένου για την αξιοπιστία και την ειλικρίνεια της δήλωσης. Ο πάροχος περιεχομένου θα πρέπει, ως εκ τούτου, κατά κανόνα, να κληθεί ως μάρτυρας κατά τη διάρκεια της δίκης.

Κεφάλαιο 9
Ιδιωτικότητα στα Social Networks

9.1 Ιδιωτικότητα

Η ιδιωτικότητα από κάποιους ορίζεται ως ο τρόπος ελέγχου των προσωπικών μας δεδομένων, ενώ από κάποιους άλλους [Zhang, 2014:1890] είναι ο προσδιορισμός του ποσοστού που κοινοποιούνται οι προσωπικές τους πληροφορίες. Η προστασία της ιδιωτικότητα δεν θεωρείται η προστασία μιας ιδιοκτησίας ή ενός φυσικού προσώπου σύμφωνα με τα ανθρώπινα δικαιώματα (άρθρο 8/1950), αλλά κάτι πολύ πιο πολύπλοκο όπου υπάρχει η ανάμιξη της τεχνολογίας και της επιστήμης της πληροφορίας.

Στην σημερινή εποχή με την συνεχή χρήση του διαδικτύου η προστασίας της ιδιωτικότητας αποτελεί θέμα συνεχούς μελέτης, διότι οι διαχειριστές όλων των υπαρχόντων πληροφοριακών συστημάτων, όπου γίνεται αποθήκευση των προσωπικών δεδομένων, μπορούν εύκολα πλέον να τα αντιγράψουν αλλά και να μεταφέρουν χωρίς να έχουν πλέον οι χρήστες κανένα έλεγχο στις προσωπικές τους πληροφορίες. Το θεσμικό πλαίσιο προσφέρει προστασία των δεδομένων του χρήστη μέσω της επίγνωσης και του περιορισμού του τρόπου χρήσης τους (συναίνεση του χρήστη).

Στις σύγχρονες δημοκρατικές κοινωνίες η προστασία της ιδιωτικότητας είναι απαραίτητη για να μην περιορίζονται οι ελευθερίες και τα ανθρώπινα δικαιώματα, από την άλλη η κοινωνία δείχνει ανησυχία αναφορικά με την ιδιωτικότητα, διότι υπάρχει έλλειψη εμπιστοσύνης. Η ύπαρξη της εμπιστοσύνης βοηθάει ώστε να υπάρχει βιώσιμη ανάπτυξη και κοινωνική σταθερότητα δηλαδή απαιτούνται αξιόπιστα συστήματα όπου

ενσωματώνονται οι αρχές και οι μηχανισμοί προστασίας της ιδιωτικότητας και των προσωπικών δεδομένων.

Για να μπορεί ένα σύστημα να διατηρεί την ιδιωτικότητα του χρήστη όταν αυτός το χρησιμοποιεί, πρέπει να προσφέρει τα παρακάτω χαρακτηριστικά [Zhang, 2014:1891]:

- Ανωνυμία
- Χωρίς συνδεσιμότητα
- Χωρίς παρατηρησιμότητα
- Χρήση ελαχίστων δεδομένων

Ανωνυμία

Παλαιότερα η ζωή των ανθρώπων καθώς και οι καθημερινές τους συνήθειες γίνονταν με περισσότερη ανωνυμία αναφορικά με τα σημερινά δεδομένα [Ramim, 2006:203]. Στην σημερινή εποχή η χρήση πιστωτικών καρτών, η χρήση κινητών τηλεφώνων, αλλά και η αυξανόμενη χρήση του internet ως μέσο επικοινωνίας και συναλλαγών αποτελούν εύκολο τρόπο για τον εντοπισμό της θέσης μας.

Η έννοια της ανωνυμίας σύμφωνα με τους Pfitzmann και Hansen, [Hansen, 2011 :64] θεωρείται η κατάσταση που είναι αδύνατο να προσδιοριστεί ένα άτομο, γιατί είναι κρυμμένο μέσα σε μια ομάδα ατόμων, όπου δεν μπορεί να γίνει προσδιορισμός ποιά άτομα επικοινωνούν μεταξύ τους. Η πιθανότητα αναγνώρισης ενός ατόμου μέσα από το σύνολο εξαρτάται από το πλήθος των ατόμων, ενώ τα άτομα που απαρτίζουν το σύνολο θα πρέπει να έχουν παρόμοια συμπεριφορά. Έτσι, σύμφωνα με τα παραπάνω εφόσον ο χρήστης διασφαλίσει την ανωνυμία του μπορεί στην συνέχεια να επικοινωνήσει με κάποιον άλλον χρήστη χωρίς να αποκαλύψει την ταυτότητα του. Η ανωνυμία ενός χρήστη, αποτελεί την βασική προϋπόθεση για επίτευξη και άλλων στόχων που ενισχύουν περαιτέρω την ιδιωτικότητα και την ασφάλεια όπως είναι η μη συνδεσιμότητα και η μη παρατηρησιμότητα [Zhang, 2014:1890, Ramim, 2006:203, Hansen, 2011 :64].

Χωρίς συνδεσιμότητα

Η ύπαρξη μη συνδεσιμότητας κατά την διάρκεια μιας συναλλαγής ή επικοινωνίας, καταδεικνύει ότι σε περίπτωση που υπάρχει κάποιος που παρατηρεί τις κινήσεις και το σύνολο μιας επικοινωνίας, παρόλα αυτά δεν μπορεί να συσχετίσει συναλλαγές με αντίστοιχα άτομα. Η απαίτηση της μη-συνδεσιμότητας προστατεύει την ιδιωτικότητα των χρηστών από εξωτερικές επιθέσεις απαγορεύοντας στους δεύτερους να συνδέσουν τμήματα σχετικών πληροφοριών μεταξύ τους, και έτσι να αποκαλυφθεί η ταυτότητα των χρηστών.

Χωρίς παρατηρησιμότητα

Η μη ύπαρξη παρατηρησιμότητας σε μια υπηρεσία προσφέρει ακόμα πιο ισχυρούς μηχανισμούς ασφαλείας, διότι προστατεύει την ιδιωτικότητα των χρηστών από πιθανές επιθέσεις απαγορεύοντας των εντοπισμό των ιχνών των χρηστών κατά την διάρκεια περιήγησης τους στο internet.

Χρήση ελαχίστων δεδομένων

Κατά την αποκάλυψη των ελαχίστων δεδομένων από τον χρήστη, αποκαλύπτονται μόνο όσα στοιχεία είναι απαραίτητα για την διεκπεραίωση της εκάστοτε συναλλαγής. Στο σύστημα που γίνεται εφαρμογή αυτής της αρχής από τον χρήστη απαιτείται να παρουσιάζει μόνο εκείνες τις πληροφορίες που είναι αναγκαίες και γνωρίζει πως μόνο αυτές πρόκειται να χρησιμοποιηθούν.

9.2 Βασικά είδη ταυτοποίησης

Τα βασικά είδη ταυτοποίησης είναι τα παρακάτω:

Αυθεντικοποίηση (Authentication)

Η διαδικασία αυθεντικοποίησης πραγματοποιείται με την επιβεβαίωση της ταυτότητας ενός χρήστη. Η αυθεντικοποίηση, σε ιδιωτικά και δημόσια δίκτυα, γίνεται συνήθως με τη

χρήση κωδικών πρόσβασης, διότι αποτελεί περισσότερο απαίτηση ασφάλειας, παρά ιδιωτικότητας ενός συστήματος.

Εξουσιοδότηση (Authorization)

Η εξουσιοδότηση αποτελεί εκείνη την διαδικασία όπου παραχωρούνται τα δικαιώματα όπως εκείνο της πρόσβασης ενός χρήστη σε μια μεμονωμένη υπηρεσία ή σε πόρους πληροφορικού συστήματος. Σε περιπτώσεις όπου υπάρχει πληθώρα χρηστών σε ένα σύστημα, ο διαχειριστής εξουσιοδοτεί τον κάθε χρήστη με τα αντίστοιχα δικαιώματα, σύμφωνα με το ρόλο τους και τις υποχρεώσεις τους στο σύστημα.

9.3 Τεχνολογίες Προστασίας της Ιδιωτικότητας (PETs)

Οι πιο διαδεδομένες τεχνολογίες που επιτρέπουν την ανώνυμη και μη ανιχνεύσιμη επικοινωνία σε ένα δίκτυο ενσωματώνουν αρχές όπως:

- Μείωση σε όσο το δυνατόν μεγαλύτερο ποσοστό του κίνδυνου παραβίασης των αρχών της ιδιωτικής ζωής και νομοθεσίας.
- Μείωση στο ελάχιστο της ποσότητας των μετα-δεδομένων που συλλέγονται για τους χρήστες.
- Ενθάρρυνση των χρηστών αναφορικά με την διατήρηση και έλεγχο των προσωπικών τους δεδομένων.

Υπάρχουν πολλά διαθέσιμα εργαλεία όπως browser add-ons, tools και apps τα οποία ειδικεύονται στην προστασία της ιδιωτικότητας του χρήστη, ενώ οι εξυπηρετητές χρησιμοποιούν τεχνικές κρυπτογράφησης της βάσης δεδομένων και της επικοινωνίας με πρωτόκολλα όπως το SSL. Όλες οι παραπάνω τεχνολογίες έχουν συμπληρωματική δράση και σχετίζονται πιο πολύ με την ασφάλεια των δεδομένων. Η μετάβαση στο web 2.0 και η συνεχής ενημέρωση των χρηστών του αναφορικά με την ιδιωτικότητα, έχουν προσδώσει μεγαλύτερη βαρύτητα στην έννοια της εμπιστοσύνης στο διαδίκτυο [Peng, 2014:298,Ma, 2013:727].

Πιστοποιητικό

Με τον όρο πιστοποιητικό θεωρείται οποιοδήποτε μέσο χρησιμοποιείται για την απόδειξη της εγκυρότητας των στοιχείων του κατόχου τους σε έναν άλλον χρήστη, είτε πρόκειται για ταυτότητα, είτε χαρακτηριστικό, στα πλαίσια πχ. ενός αιτήματος πρόσβασης σε μια υπηρεσία. Η έκδοση του πιστοποιητικό γίνεται από αρμόδια αρχή και ένα παράδειγμα τέτοιου πιστοποιητικού είναι η άδεια οδήγησης της οποίας ο κάτοχος τη δείχνει σε κάποια αρχή ελέγχου (αστυνομία) για να αποδείξει ότι δικαιούται να οδηγεί κάποιο όχημα.

Ψηφιακά Πιστοποιητικά (Digital Credentials)

Η χρήση ψηφιακών πιστοποιητικών (digital certificates ή credentials) αυξάνεται διαρκώς, διότι όλο και περισσότερες χώρες υιοθετούν ψηφιακές ταυτότητες για χρήση σε συναλλαγές. Η έκδοση πραγματοποιείται όταν ο χρήστης επιβεβαιώσει την αρμόδια αρχή (πάροχος) ότι κατέχει ένα συγκεκριμένο σύνολο ιδιοτήτων ή χαρακτηριστικών και ο πάροχος με τη σειρά του να επιβεβαιώσει ότι αυτό είναι αληθές.

Ανώνυμο Πιστοποιητικό

Το κλασσικό πιστοποιητικό όταν ο χρήστης το παρουσιάσει σε κάποιο τρίτο όπως π.χ. ένας χρήσης επιδεικνύει το δίπλωμα οδήγησης σε μια εταιρεία ενοικιάσεως αυτοκινήτων, έχει δύο βασικά μειονεκτήματα, το πρώτο είναι ότι ο χρήστης απευθείας διαθέτει όλα τα στοιχεία του χωρίς να είναι απαραίτητο και το δεύτερο είναι ότι αυτός που κάνει την επιβεβαίωση μπορεί να χρησιμοποιήσει τα προσωπικά στοιχεία για να εξαπατήσει το χρήστη. Η χρήση των ανώνυμων πιστοποιητικών αντιμετωπίζει τα παραπάνω δύο προβλήματα. Με το ανώνυμο πιστοποιητικό, δεν υπάρχει άμεση μεταβίβαση του πιστοποιητικού στην αρμόδια αρχή επιβεβαίωσης, επίσης ο χρήστης κάνει αποστολή μόνο την κρυπτογραφημένη απόδειξη ότι κατέχει το απαραίτητο πιστοποιητικό και επιλέγει τα συγκεκριμένα στοιχεία που θα αποστείλει ώστε να πληρούν το δεδομένο αίτημα, τα υπόλοιπα στοιχεία παραμένουν κρυφά [Neven, 2008].

Οι τεχνολογίες διαχείρισης στοιχείων ταυτότητας αποκτούν όλο και μεγαλύτερη σημασία στην στη σημερινή εποχή. Ένα στοιχειώδης μοντέλο μιας διαδικασίας ταυτοποίησης περιλαμβάνει:

- χρήστες (users)
- πάροχο της ταυτότητας (identity provider)
- πάροχο της υπηρεσίας (service provider)

Ο κάθε χρήστης που θέλει να αποκτήσει πρόσβαση σε μια υπηρεσία ακολουθεί την παρακάτω διαδικασία, αρχικά ο identity provider εκτελεί τις λειτουργίες της αυθεντικοποίησης και εξουσιοδότησης, στην συνέχεια μόλις ο χρήστης λάβει την εξουσιοδότηση τότε μόνο ο πάροχος της υπηρεσίας του δίνει το δικαίωμα πρόσβασης.

Το πως χρησιμοποιούνται οι τεχνολογίες ταυτοποίησης εξαρτάται από το ποιος τις χρησιμοποιεί. Οι κατηγορίες εμπλεκόμενων συνήθως είναι:

- χρήστες
- επιχειρήσεις
- υπηρεσίες
- κυβερνητικοί οργανισμοί

Οι χρήστες απαιτούν να αποκτούν πρόσβαση με τρόπο εύχρηστο και ασφαλή. Παρόλα αυτά οι περισσότεροι χρήστες σήμερα χρησιμοποιούν πλήθος ζευγών username-password που χρησιμοποιούν για πρόσβαση σε διάφορες “online” υπηρεσίες. Επειδή η απομνημόνευση όλων των πιστοποιητικών είναι πολύ δύσκολη, συχνά οι χρήστες χρησιμοποιούν τα ίδια username και password και συνήθως αδύναμους και μικρούς σε μήκος συνδυασμούς [Yuan, 2013:637].

Από την άλλη μεριά ο χρήστης δεν έχει καθόλου έλεγχο στα προσωπικά του δεδομένα, καθώς δεν γνωρίζει που βρίσκονται, ποιος τα χρησιμοποιεί και για ποιο σκοπό. Ιδιαίτερα στην χώρα μας, αναφορικά με θέματα προστασίας δεδομένων υπάρχει μια αναντιστοιχία κανονιστικής ρύθμισης και τεχνολογικών εξελίξεων. Στις υπόλοιπες Ευρωπαϊκές χώρες όπου υπάρχει σχετική νομοθεσία και ακολουθούν κατά πόδας τις τεχνολογικές εξελίξεις, στη πραγματικότητα δεν υπάρχει ενεργός έλεγχος σε παρόχους και οργανισμούς, π.χ. δεν υπάρχει όριο στην πληροφορία που μπορεί να καταχωρηθεί, στην ανάλυση που μπορεί

να γίνει ούτε στο χρονικό διάστημα για το οποίο μπορεί να τηρηθεί [Masoumzadeh, 2012:881].

Στις επιχειρήσεις το ενδιαφέρον τους επικεντρώνεται σε συστήματα ταυτοποίησης για λόγους ασφάλειας και αποδοτικότητας των συναλλαγών. Έτσι, εκτός από το να παρέχουν προϊόντα και υπηρεσίες, προσπαθούν να αναπτύξουν σχέσεις με τους πελάτες τους ενισχύοντας την εμπιστοσύνη τους. Η επικοινωνία μεταξύ πελάτη και επιχείρησης τις περισσότερες φορές πραγματοποιείται με ψηφιακό τρόπο μέσω εισαγωγής username και password στο website της εταιρείας.

Στο θέμα της απόδοσης η διαχείριση των προσωπικών πληροφοριών με κατάλληλο τρόπο αυξάνει την αποδοτικότητα δηλαδή την ταχύτητα και την ακρίβεια διεκπεραίωσης των συναλλαγών. Ταυτόχρονα όμως όπως αναφέρθηκε εγείρονται ζητήματα που αφορούν την ιδιωτικότητα του χρήστη-πελάτη.

Όσο αναφορά τους κυβερνητικούς οργανισμούς το ενδιαφέρον για τις τεχνολογίες ταυτοποίησης είναι έντονο διότι η κυβέρνηση αποτελεί έναν από τους μεγαλύτερους παρόχους πολλών εγγράφων ταυτότητας όπως αστυνομικές ταυτότητες, διαβατήρια, άδειες οδήγησης. Γι' αυτό το λόγο συντάσσονται νόμοι που αφορούν τη συλλογή προσωπικών δεδομένων και ρυθμιστικά πλαίσια.

9.4 Σύγχρονα Συστήματα Ταυτοποίησης

Στην σημερινή εποχή δίνεται μεγαλύτερη έμφαση στα κεντρικοποιημένα συστήματα ταυτοποίησης (Federated Identity Management). Η διαδικασία της ταυτοποίησης περιλαμβάνει διάφορα domains τα οποία μπορεί να ανήκουν είτε στην ίδια οντότητα είτε σε άλλο οργανισμό. Η ψηφιακή ταυτότητα και τα χαρακτηριστικά (attributes) του χρήστη ουσιαστικά βρίσκονται αποθηκευμένα σε πολλαπλά συστήματα. Με αυτόν τον τρόπο, με ένα συνθηματικό, (username- password), δηλαδή με ένα single sign-on (SSO), ο χρήστης έχει πρόσβαση σε πολλαπλά συστήματα.

Ψηφιακές πλατφόρμες που επιτρέπουν στους χρήστες να συνδεθούν διαμέσου τρίτων (third-party) websites, εφαρμογών, συσκευών κινητής τηλεφωνίας με την ήδη υπάρχουσα ταυτότητα τους ,πχ enable social login, είναι: το Facebook, Google, Yahoo!, Twitter, LinkedIn και PayPal. Επίσης προσφέρεται και η δυνατότητα μεταφοράς πληροφοριών από το προφίλ του χρήστη σε άλλα site όπως όνομα , διεύθυνση, email κτλ.

Απαραίτητη προϋπόθεση για την ανάπτυξη των κεντρικοποιημένων συστημάτων είναι η ύπαρξη εμπιστοσύνης και επικοινωνίας μεταξύ των οργανισμών, μιας και τα πλεονεκτήματα είναι αρκετά:

- Μείωση του κόστους αναβάθμισης και αύξηση της ασφάλειας
- Η επαλήθευση τη ταυτότητας του χρήστη πραγματοποιείται μόνο τη πρώτη φορά και μπορεί να χρησιμοποιηθεί για πρόσβαση σε πολλά σημεία.
- Ο χρήστης είναι υπεύθυνος για τα στοιχεία που διαμοιράζεται στο βαθμό που αυτός επιθυμεί.

Λόγω των παραπάνω χαρακτηριστικών, πλεονεκτημάτων τα κεντρικοποιημένα συστήματα είναι ιδιαίτερα δημοφιλή για την αποδοτικότητα τους, αφού οι χρήστες πλέον δεν χρειάζεται να δημιουργούν καινούργιους λογαριασμούς συνεχώς, ούτε να κάνουν αλλεπάλληλα login σε διαφορετικές υπηρεσίες.

9.5 Πρακτικοί Τρόποι Προστασίας Ιδιωτικότητας στο Twitter και Instagram

Η ιδιωτική ζωή και τα προσωπικά μας δεδομένα, είναι πολύτιμα. Για τον καθένα από εμάς υπάρχουν ευαίσθητες πληροφορίες που δεν θα θέλαμε να μοιραστούμε δημόσια. Ως προσωπικά δεδομένα εννοούμε τις πληροφορίες που αναφέρονται στο πρόσωπό μας, όπως το ονοματεπώνυμο, η διεύθυνση, ο αριθμός του κινητού τηλεφώνου, τα μέρη όπου ταξιδεύουμε ή βρισκόμαστε, με το τελευταίο να αποτελεί και το αντικείμενο της παρούσας διπλωματικής εργασίας. Διατηρώντας τον έλεγχο των προσωπικών μας δεδομένων, διατηρείς και τον έλεγχο της ιδιωτικής μας ζωής. Από τα παραπάνω

συμπεραίνουμε ότι η πρώτη γραμμή άμυνας εναντίον του κακόβουλου χρήστη είναι η ευαισθητοποίηση σε θέματα ασφάλειας και διαχείρισης προσωπικών δεδομένων.

Ειδικά για το Twitter και σε ότι έχει να κάνει με τον διαμοιρασμό της γεωγραφικής μας τοποθεσίας, μπορούμε να την ενεργοποιήσουμε ειδικά για ένα μήνυμα ή να είναι γενικά ενεργή για όλα τα μηνύματά μας. Εξ' ορισμού είναι απενεργοποιημένη η συγκεκριμένη λειτουργία και θα πρέπει να την ενεργοποιήσουμε συνειδητά. Όπως είδαμε σε προηγούμενα κεφάλαια και μέσα από το πρόγραμμά μας, οι ετικέτες με γεωγραφικά δεδομένα είναι δύο: location (δηλώνουμε πόλη ή περιοχή) και precise gps location (ακριβής γεωγραφικές συντεταγμένες από τον δέκτη GPS του smartphone μας). Θα πρέπει να είμαστε ιδιαίτερα επιφυλακτικοί, εάν χρησιμοποιούμε μια παλαιότερη έκδοση Twitter για IOS ή Twitter για Android, καθότι οι αρχικές ρυθμίσεις μπορεί να είναι διαφορετικές και να αποκαλύπτουν στοιχεία για την γεωγραφική μας θέση. Επιπλέον θα πρέπει να είμαστε προσεκτικοί στους όρους χρήσης εφαρμογών τρίτων κατασκευαστών που χρησιμοποιούν το Twitter API, καθότι ενδέχεται να αποκαλύπτουν την γεωγραφική μας θέση, χωρίς να είναι εξαρχής ξεκάθαρο. Θα πρέπει να αναγνωρίσουμε στο Twitter ότι προσφέρει την δυνατότητα να αφαιρέσουμε εμείς οι ίδιοι σαν τελικοί χρήστες, τα γεωγραφικά δεδομένα από όλα τα μηνύματα μας συνολικά [Twitter, 2014].

Σχετικά με το Instagram, οι ρυθμίσεις που αφορούν την εισαγωγή ή όχι γεωγραφικών συντεταγμένων, βρίσκονται κάτω από τις γενικές ρυθμίσεις του smartphone. Το Instagram κάνει πολύ εύκολο τον διαμοιρασμό των φωτογραφιών μέσω του Facebook και στην συγκεκριμένη περίπτωση θα πρέπει να είμαστε πολύ προσεκτικοί με την αλληλεπίδραση των δύο αυτών πλατφορμών κοινωνικής δικτύωσης. Είμαστε ιδιαίτερα προσεκτικοί όσον αφορά φωτογραφίες που μοιραζόμαστε δημόσια στο Instagram και σχετίζονται με τον τόπο κατοικίας μας ή την εργασία μας.

Κεφάλαιο 10
Συμπεράσματα

Το βασικό ερευνητικό ερώτημα που υποστηρίχθηκε στην παρούσα διπλωματική εργασία, είχε να κάνει με το κατά πόσο είναι δυνατόν κάποιος κακόβουλος χρήστης να παρακολουθεί, χωρίς την συγκατάθεση του χρήστη στόχου, την δραστηριότητα του στα δύο μεγάλα μέσα κοινωνικής δικτύωσης Twitter και Instagram και να εξάγει συμπεράσματα, δεδομένα που αφορούν την τοποθεσία του ανά χρονικό σημείο.

Ακολουθώντας τεχνικές OSINT και με τις μεθόδους που αναλύθηκαν σε αντίστοιχα κεφάλαια καταφέραμε να αποκτήσουμε πρόσβαση σε Twitter, Instagram, Google APIs και cloud vm, χωρίς να αποκαλύψουμε την πραγματική μας ταυτότητα και σκοπούς και χωρίς να μπορεί κάτι να συνδεθεί πίσω σε εμάς τουλάχιστον εύκολα. Οι τεχνικές αυτές λειτουργούσαν όπως περιγράφονται κατά το χρονικό διάστημα εκπόνησης της συγκεκριμένης εργασίας δηλαδή από τον Ιούλιο έως τον Νοέμβριο του 2015.

Η σύνθεση των παραπάνω λειτουργιών, έγινε μέσω του SocialMap python script. Τα αποτελέσματα του, με βάση πραγματικούς στόχους που αντιστοιχούν σε λογαριασμούς διάσημων προσωπικοτήτων από τον Ελληνικό χώρο, καταδεικνύουν ότι τα αποτυπώματα που αφήνουν πίσω τους οι δραστηριότητές μας στα δύο εν λόγω κοινωνικά δίκτυα, μπορούν να αναγνωστούν και να χρησιμοποιηθούν από τρίτους, χωρίς την συγκατάθεσή και την γνώση του χρήστη στόχου. Επικεντρωθήκαμε μόνο στην αποτύπωση του γεωγραφικού στίγματος του χρήστη, αποδεικνύοντας ότι η συνεχής παρακολούθηση ενός πολίτη όπως περιγράφεται μέσα στο μυθιστόρημα του George Orwell “1984” και την δράση του χαρακτήρα του Μεγάλου Αδερφού (Big Brother is watching you), μπορεί να γίνει πραγματικότητα.

Μπορούμε να προστατευτούμε πλήρως σαν τελικοί χρήστες από τέτοιες κακόβουλες ενέργειες; Μάλλον όχι 100%, γιατί η ίδια η φύση των μέσων κοινωνικής δικτύωσης, απαιτεί την δημιουργία περιεχομένου από τον ίδιο τον χρήστη, που δυστυχώς τείνει να μεταφέρει την προσωπικότητά του στο μέσο. Η χρήση της κοινής λογικής, ο διαχωρισμός και η διαφύλαξη των ευαίσθητων προσωπικών πληροφοριών και η ευαισθητοποίηση

πάνω σε θέματα ασφάλειας (security awareness), μπορούν να βοηθήσουν σημαντικά στον περιορισμό της έκθεσής μας σε τέτοιου είδους επιθέσεις.

Στις προτάσεις βελτιώσεων της συγκεκριμένης εργασίας μελλοντικά, θα μπορούσαμε να αναφέρουμε την υποστήριξη περισσότερων πλατφορμών κοινωνικής δικτύωσης και την δημιουργία ενός γραφικού web interface / εφαρμογής με φιλικό και εύχρηστο προς τον χρήστη περιβάλλον λειτουργίας. Επιπλέον θα μπορούσαν να ενσωματωθούν λειτουργίες δημιουργίας αυτοματοποιημένων εκθέσεων (reports) ανά χρήστη στόχο, για δοσμένο συγκεκριμένο χρονικό διάστημα, και ο αντίστοιχος βαθμός έκθεσης του. Μια ακόμα δυνατότητα – επέκταση που θα μπορούσε να ενσωματωθεί είναι η πρόβλεψη της μελλοντικής θέσης του στόχου με βάση το προηγούμενο ιστορικό του και την συχνότητα επισκέψεων σε συγκεκριμένη τοποθεσία. Τέλος θα μπορούσε η συγκεκριμένη εργασία, να προσαρμοστεί σε περιπτώσεις μεγάλων εταιριών, οι οποίες θα ήθελαν να προσδιορίσουν τον βαθμό απειλής διαρροών σημαντικών πληροφοριών από τους ίδιους τους υπαλλήλους τους (insider threats) ή να αποτυπώσουν το ψυχολογικό – κοινωνικό προφίλ υπαλλήλων τους σε κρίσιμες θέσεις, όπως αυτό συμπεραίνεται από την παρουσία τους στα social media.

Παράρτημα Α
Python Script SocialMap

Παρακάτω ακολουθεί ο κώδικας με σχόλια από το εν λόγω **python script SocialMap.py**.

```
001 import tweepy
002 from instagram.client import InstagramAPI
003 import csv
004 import requests
005 import json
006 import codecs
007 from geopy.geocoders import GoogleV3
008 requests.packages.urllib3.disable_warnings()
009
010 #Twitter API credentials
011 consumer_key = "---------------"
012 consumer_secret = "---------------"
013 access_key = "---------------"
014 access_secret = "---------------"
015
016 #this function uses the geocode Google api to get gps location from
the given location name
017 def geocode(loc):
018     print "geocoding %s it might take a while ...." % loc
019     try:
```

```
020         #check if it can be coded
021         geolocator = GoogleV3(api_key='---------------')
022         location = geolocator.geocode(loc)
023         if location:
024             #return if it can be
025             return [location.latitude,location.longitude]
026     except:
027         pass
028 
029     return None
030 
031 
032 #this function generates the map
033 def generate_map(username,gps):
034     print "Generating map"
035     #list of all marker data to be put
036     gpsData = []
037     #merging the two different data sets
038     for dataset in gps:
039         for i,data in enumerate(dataset[1]):
040             #if it's the very first, we put the green marker on it
041             if i == 0:
042 
data.append("http://chart.apis.google.com/chart?chst=d_map_pin_letter&chld=
1|008A00|FFFFFF")
043             else:
044                 #if it's a tweet, blue
```

```
045                 if(dataset[0] == 0):
046                     color =
"http://chart.apis.google.com/chart?chst=d_map_pin_letter&chld=%s|094AB2|FF
FFFF" % (i+1)
047                 #if it's a insta, red
048                 else:
049                     color =
"http://chart.apis.google.com/chart?chst=d_map_pin_letter&chld=%s|D24726|FF
FFFF" % (i+1)
050                 data.append(color)
051
052             gpsData.append(data)
053
054     #print gpsData
055
056     if(len(gpsData) > 0):
057         #for a map string with the necessary html in it
058         cordString = ''
059         for cord in gpsData:
060             cordString += '["%s",{ lat: %s, lng: %s }, "%s"],' %
(cord[2],cord[1],cord[0],cord[3])
061         html = '''
062                     <!doctype html>
063                     <html>
064                     <head>
065                     <title>Maps</title>
066                     <meta name="viewport" content="width=device-width,
user-scalable=no">
067                     <style>
```

```
068                     html,body{
069                         height:100%%;
070                         padding:0px;
071                         margin:0px;
072                     }
073                     #map-canvas{
074                         height:100%%;
075                     }
076                     </style>
077
078                     <script type="text/javascript"
src="https://maps.googleapis.com/maps/api/js?key=---------------
&sensor=false"></script>
079                     <script type="text/javascript">
080                     function initialize() {
081                             //stores the marker data
082                             var myLatLng = [%s];
083                             //initialize the map
084                               window.map = new
google.maps.Map(document.getElementById('map-canvas'), {
085                                 zoom: 8,
086                                 center: myLatLng[0][1]
087                               });
088
089                             var LIMIT = 100;
090                             var maxlen = (myLatLng.length > 100) ? 100
: myLatLng.length;
091
```

```
092                             //sort the array by date
093                             myLatLng = sortArray(myLatLng);
094
095                             //this holds all the markers we create
096                             var markerset = [];
097
098                             //loop all the marker data
099                             for(var i=0;i<maxlen;i++){
100                                 var found = false;
101                                 //first check if the marker is already
positioned in current gps location
102                                 for(var j=0;j<markerset.length;j++){
103                                     //if positioned we simply add the
time data to existing marker
104                                     if(markerset[j].lat ==
myLatLng[i][1].lat && markerset[j].lng == myLatLng[i][1].lng){
105                                         var marker = markerset[j];
106                                         marker.info.content =
marker.info.content + "<br />" + myLatLng[i][0];
107                                         //markerset.push(marker);
108                                         found = true;
109                                         break;
110                                     }
111                                 }
112
113                               //if not a new marker is created
114                               if(!found){
```

```
115                                     //again a crude checking for green
pin, this has to be done since merging is done
116                                     if(myLatLng[i][2].indexOf("008A00") >
-1){
117                                        var pinIcon = new
google.maps.MarkerImage(
118                                          myLatLng[i][2],
119                                          null, /* size is determined at
runtime */
120                                          null, /* origin is 0,0 */
121                                          null, /* anchor is bottom
center of the scaled image */
122                                          new google.maps.Size(40, 61)
123                                        );
124                                     }else{
125                                        var pinIcon = new
google.maps.MarkerImage(
126                                          myLatLng[i][2],
127                                          null, /* size is determined at
runtime */
128                                          null, /* origin is 0,0 */
129                                          null, /* anchor is bottom
center of the scaled image */
130                                          new google.maps.Size(21, 34)
131                                        );
132                                     }
133
134                                     //add new marker
135                                     var marker = new google.maps.Marker({
```

```
136                                 position: myLatLng[i][1],
137                                 map: map,
138                                 icon: pinIcon
139                               });
140
141                               //store the lat,lng to use in the
check-for-duplicates code
142                               marker.lat = myLatLng[i][1].lat;
143                               marker.lng = myLatLng[i][1].lng;
144
145                               //adding the info window
146                               marker.info = new
google.maps.InfoWindow({
147                                 content: myLatLng[i][0]
148                               });
149
150                               //adding the listener to show the
window on click
151                               google.maps.event.addListener(marker,
'click', function() {
152                                   this.info.open(map, this);
153                                });
154
155                                //add the marker to the set
156                                markerset.push(marker);
157                             }
158                         }
159
```

```
160                             //drawing all the markers with .5s interval
161                             var pointer = markerset.length - 1;
162                             window.drawing = setInterval(function(){
163 draw(markerset[pointer],markerset[pointer-1]);
164                                 pointer--;
165                                 if(pointer == 0){
166                                     clearInterval(drawing);
167                                 }
168                             },500);
169
170
171                     }
172
173                     //function to draw the arrow
174                     function draw(from,to){
175                         var lineSymbol = {
176                             path:
google.maps.SymbolPath.FORWARD_CLOSED_ARROW
177                         };
178
179                         // Create the polyline and add the symbol via
the 'icons' property.
180                         var line = new google.maps.Polyline({
181                             path: [{lat: from.lat, lng: from.lng},
{lat: to.lat, lng: to.lng}],
182                             icons: [{
183                                 icon: lineSymbol,
```

```
184                                 offset: '100%%'
185                             }],
186                             map: map
187                         });
188                     map.panTo({lat: to.lat, lng: to.lng});
189                 }
190 
191                 //function to sort the merged list with respective
to the date-time
192                 function sortArray(ary){
193                     ary.sort(function(x,y){
194                         var xd = Date.parse(x[0]);
195                         var yd = Date.parse(y[0]);
196                         if(xd < yd) return 1;
197                         if(xd > yd) return -1;
198                         return 0;
199                     });
200                     return ary;
201                 }
202 
203                         google.maps.event.addDomListener(window,
'load', initialize);
204                     </script>
205                 </head>
206                 <body>
207 
208                 <div id="map-canvas">
```

```
209                     </div>
210
211                     </body>
212                     </html>
213
214                 ''' % (cordString)
215         #write it to the html file
216         fhandle = open('%s_map.html' % username, 'w')
217         fhandle.write(html)
218         fhandle.close()
219     else:
220         print "NO GPS DATA"
221
222 #function to run the main process
223 def get_tweets(username,isMapped,max_num = 1000):
224     print "processing....."
225     #initialize the twitter API with necessary authentications
226     auth = tweepy.OAuthHandler(consumer_key, consumer_secret)
227     auth.set_access_token(access_key, access_secret)
228     api = tweepy.API(auth)
229
230     #this contains all the tweets taken from a handle
231     alltweets = []
232     #temprorary store for tweets, to start the process we need to first
take max of 200 from handle
233     new_tweets = api.user_timeline(screen_name = username,count=200)
234     #then we extend our alltweets list with these
```

```
235     if(len(new_tweets) > 0):
236         alltweets.extend(new_tweets)
237         #find the last tweet we got since it's needed to get tweets
after than
238         oldest = alltweets[-1].id - 1
239
240         #max num is given to stop the loop from grabbing, maximum
amount for max_num is 1000
241         while (len(new_tweets) > 0 and len(alltweets) <= max_num):
242                 #again repeat the procedure above, now the max_id
parameter passed
243                 new_tweets = api.user_timeline(screen_name =
username,count=200,max_id=oldest)
244                 alltweets.extend(new_tweets)
245                 oldest = alltweets[-1].id - 1
246                 #just debug info
247                 print "...%s tweets downloaded so far" %
(len(alltweets))
248
249         #open the file for dumping the results, but we open it with
utf-8 encoding to preserve unicode
250         fhandle = codecs.open('%s_tweets.csv' % username, 'w', "utf-8")
251
252         #store the gps separtely
253         cordinateList = []
254
255         #geocode temp
256         lastGeoLocation = None
257         lastGeoCode = None
```

```
258
259         #loop the tweets
260         for tweet in alltweets:
261             #try to find gps coordinated, if found we structure it with
"space" in between
262             if tweet._json['coordinates']:
263                 tagGeo = "[%s %s]" %
(tweet._json['coordinates']['coordinates'][0],tweet._json['coordinates']['c
oordinates'][1])
264
cordinateList.append([tweet._json['coordinates']['coordinates'][0],tweet._j
son['coordinates']['coordinates'][1], tweet.created_at])
265             else:
266                 tagGeo = ''
267             #try to find location details
268             if tweet._json['user']['location']:
269                 tagLoc = "-
".join(tweet._json['user']['location'].split(','))
270             else:
271                 tagLoc = ''
272
273             if len(tagGeo) == 0 and len(tagLoc) > 0:
274                 if lastGeoLocation != None and lastGeoLocation ==
tweet._json['user']['location']:
275                     geoloc = lastGeoCode
276                 else:
277                     geoloc = geocode(tweet._json['user']['location'])
278                     lastGeoLocation = tweet._json['user']['location']
279                     lastGeoCode = geoloc
```

```
280
281                 if geoloc:
282                     tagGeo = "[%s %s]" % (geoloc[1],geoloc[0])
283
cordinateList.append([geoloc[1],geoloc[0],tweet.created_at])
284
285             #print("%s,%s,%s,%s" %
(tweet.created_at,tweet.text.encode("ascii",'ignore'),tagGeo,tagLoc))
286             #write the data to the file as a single line
287             fhandle.write("%s,%s,%s,%s \n" %
(tweet.created_at,tweet.text,tagGeo,tagLoc))
288
289         #close the file handler
290         fhandle.close()
291
292         if isMapped:
293             return (0,cordinateList)
294             #generate_map(username, cordinateList)
295     else:
296         print "ERROR: Invalid username"
297
298 #INSTAGRAM CREDENTIALS
299 client_id = "---------------"
300 client_secret = "---------------"
301
302 #function to grab the data from instagram
303 def grab_instagram_media(username, isMapped):
```

```
304     print "processing....."
305     if(len(username) == 0):
306         print "ERROR: empty username"
307         return
308
309     #connect to the api with unauthorized api call
310     api = InstagramAPI(client_id=client_id,
client_secret=client_secret)
311     #first search for a user with given username
312     search = api.user_search(q=username,count=1)
313     if search and len(search) == 1:
314         #if found get the userid, which is  used in api calls
315         userid = search[0].id
316         #all media is stored here
317         all_media = []
318         recent_media, next_ = api.user_recent_media(user_id=userid,
count=40)
319         #grab the first set of media and insert into our super list
320         all_media.extend(recent_media)
321         while (next_ and len(all_media) < 200):
322             #continue the process using the next_ parameter to get the
next set
323             recent_media, next_ = api.user_recent_media(user_id=userid,
count=40, with_next_url=next_)
324
325             all_media.extend(recent_media)
326
327             print "...%s media downloaded so far" % (len(all_media))
```

```
328
329         #open the file for dumping the results, but we open it with
utf-8 encoding to preserve unicode
330         fhandle = codecs.open('%s_instas.csv' % username, 'w', "utf-8")
331
332         #store the gps separtely
333         cordinateList = []
334         #loop all the media acquired
335         for media in all_media:
336             #check if the location attribute exist since objects are
returned
337             if hasattr(media,'location') and media.location != None and
media.location.point != None:
338                 tagGeo = "[%s %s]" %
(media.location.point.latitude,media.location.point.longitude)
339
cordinateList.append([media.location.point.longitude,media.location.point.l
atitude,media.created_time])
340             else:
341                 tagGeo = ""
342             #check for caption
343             if hasattr(media,'caption') and media.caption:
344                 if media.caption:
345                     caption = media.caption.text
346                 else:
347                     caption = ""
348             else:
349                 caption = ""
350             #write the output to the files
```

```
351             fhandle.write("%s,%s,%s \n" %
(media.created_time,caption,tagGeo))
352         #close the connection
353         fhandle.close()
354         #check if mapping is enabled
355         if isMapped:
356             #generate_map(username,cordinateList)
357             return (1,cordinateList)
358
359     else:
360         "ERROR: cant find a user or multiple users from that username"
361
362
363
364
365
366 #stuff to make the thing pretty
367 print "################################################"
368 print "Tweet/Insta Grab by K. Tzovelekis - APKY Master Thesis 2015"
369 print "################################################"
370 print "1 - Generate tweets only"
371 print "2 - Generate tweets with map"
372 print "3 - Generate instas only"
373 print "4 - Generate instas with map"
374 print "5 - Generate tweet + instas with map"
375 print "6 - EXIT"
376
```

```
377 while True:
378     inp = input("Please select an option : ")
379 
380     if inp == 6:
381         break
382     elif inp == 5:
383         try:
384             username = raw_input("Please enter twitter handle : ")
385             data1 = get_tweets(username,True)
386             username = raw_input("Please enter instagram handle : ")
387             data2 = grab_instagram_media(username,True)
388 
389             generate_map("%s_tweet_insta" % username,[data1,data2])
390             continue
391         except Exception, e:
392             pass
393 
394     username = raw_input("Please enter twitter/instagram handle : ")
395 
396 
397     try:
398         if inp == 1:
399             get_tweets(username,False)
400         elif inp == 2:
401             data = get_tweets(username,True)
402             generate_map(username,[data])
```

```
403         elif inp == 3:
404             grab_instagram_media(username, False)
405         elif inp == 4:
406             data = grab_instagram_media(username,True)
407             generate_map(username,[data])
408         else:
409             break
410     except Exception as e:
411         print e
412         print "ERROR in username or service"
413 
414     print "DONE!"
415 
```

Παράρτημα Β
Δείγματα εκτέλεσης SocialMap

Στις παρακάτω εικόνες απεικονίζονται τρεις περιπτώσεις όπου τρέξαμε το script SocialMap σε συγκεκριμένους χρήστες του Twitter και Instagram. Οι 2 πρώτες περιπτώσεις αφορούν πραγματικούς χρήστες ενώ η τρίτη περίπτωση αφορά στοιχεία που συλλέχθηκαν από έναν εικονικό χρήστη που δημιουργήθηκε για δοκιμαστικούς σκοπούς στα πλαίσια της διατριβής αυτής. Τα δεδομένα που βασίζεται η απεικόνιση σε Google Maps είναι συγκεντρωμένα σε csv αρχεία. Τα συγκεκριμένα αρχεία δεν μπορούν να δημοσιοποιηθούν αυτούσια, όπως ορίζεται στους όρους χρήσης του twitter API (II. Restrictions on Use of Licensed Materials – paragraph C - Geographic Data)[1] .

Μια πλήρης εγγραφή δεδομένων που επιστρέφει το python script από το twitter και το Instagram είναι της μορφής που φαίνεται στον παρακάτω πίνακα.

Χρονοσφραγίδα	Κείμενο twitter / Instagram	Geolocation data
2015-06-23 15:47:49	Ελλαδάρα #summertour #theatre #play #greece #ioannina #nofilter #sky	[39.665477024 20.840999688]
2015-08-20 09:36:52	Στην αγαπημένη μου Κρήτη! ♥GR My beloved Crete! #BlueMagazine #Cover #Greece #theCreteGolf #Crete #VisitGreece	[35.342567653 25.144328329]

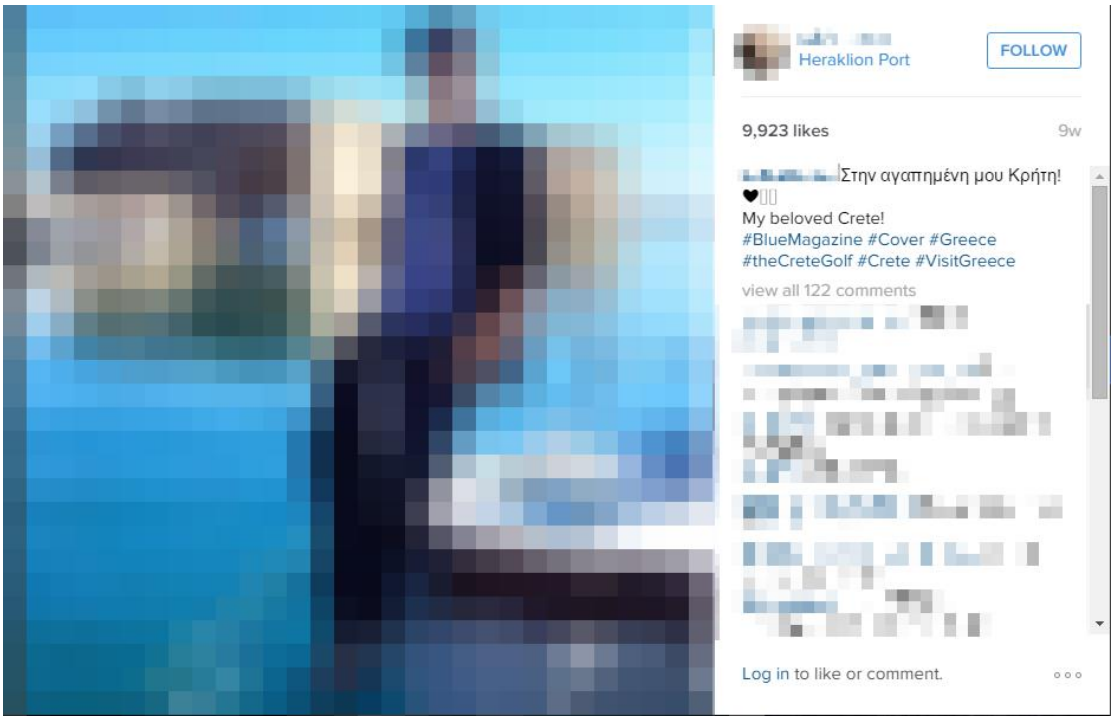
Πίνακας 3 *- Παράδειγμα twitter και Instagram εγγραφής.*

Στην παρακάτω εικόνα φαίνεται το Twitter post της παραπάνω εγγραφής 2015-06-23 15:47:49.

[1] https://dev.twitter.com/overview/terms/agreement-and-policy

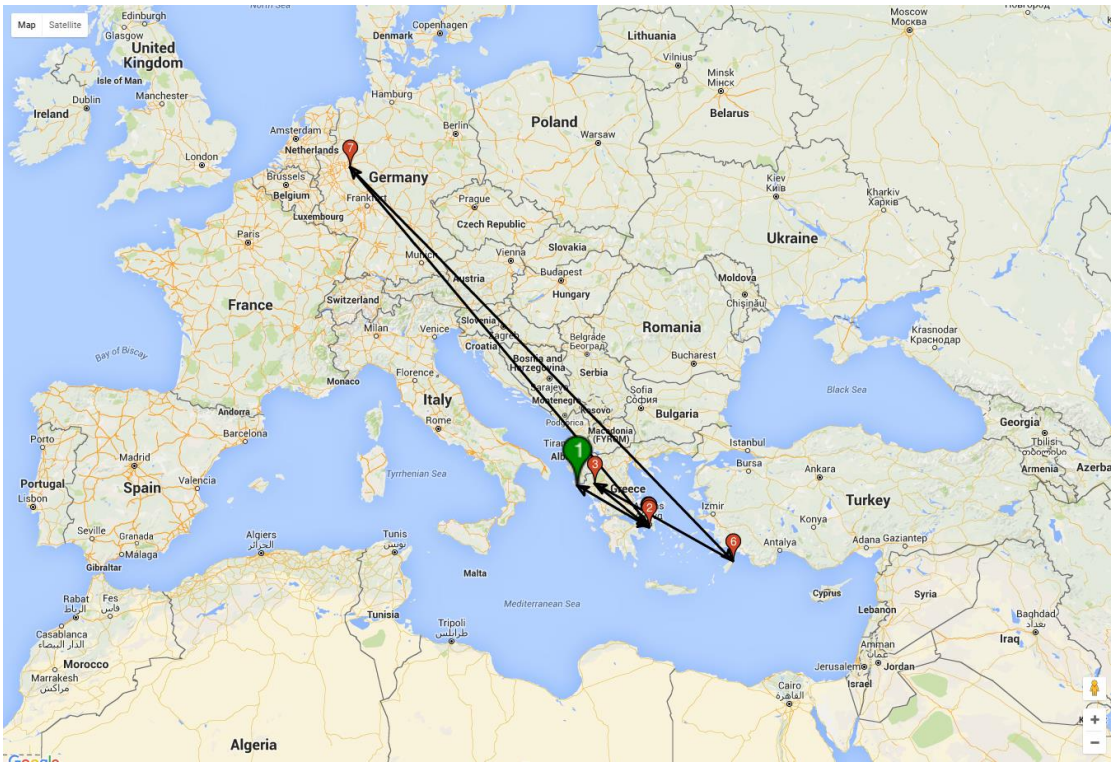
Εικόνα 42 *- Twitter post με ενσωματωμένα geolocation μετά-δεδομένα.*

Στην παρακάτω εικόνα φαίνεται το Instagram post της παραπάνω εγγραφής 2015-08-20 09:36:52.



Εικόνα 43 *- Instagram post με ενσωματωμένα geolocation μετά-δεδομένα.*

Περίπτωση 1 πραγματικού χρήστη (LN)

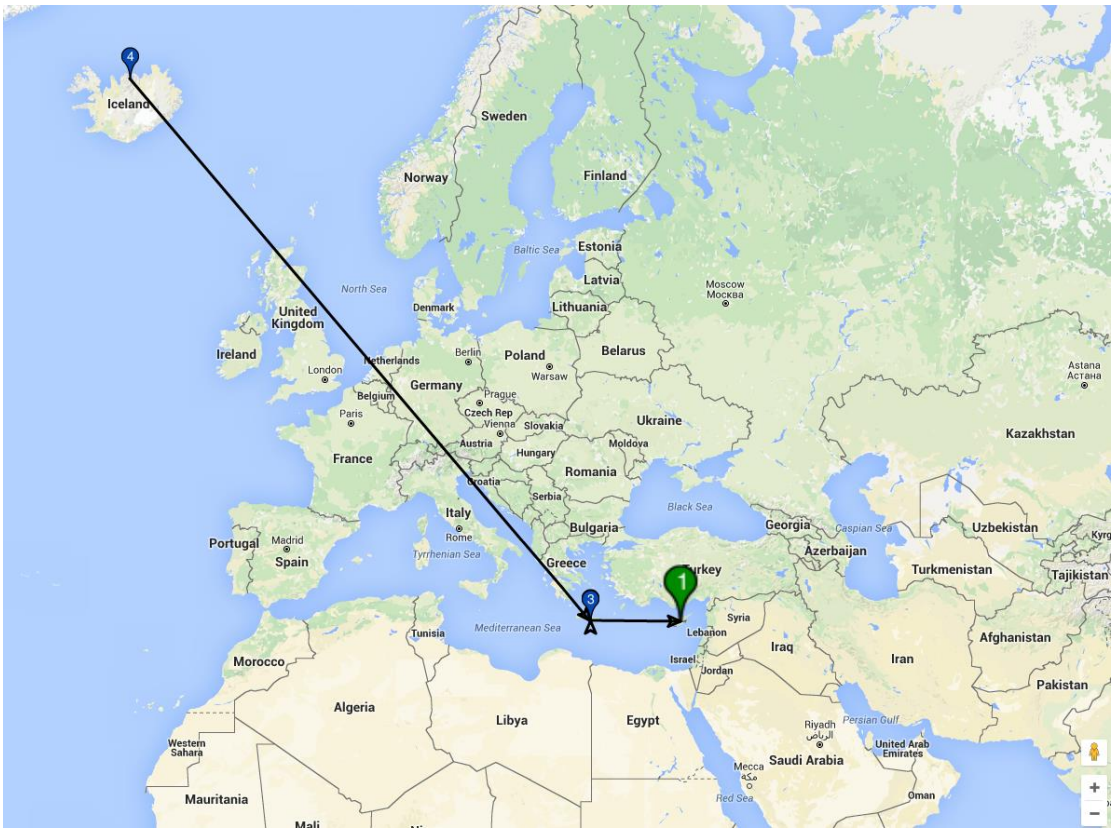


Εικόνα 44 *- Η απεικόνιση των δεδομένων που συλλέχθηκαν για τον χρήστη LR.*

Περίπτωση 2 πραγματικού χρήστη (SR)

Εικόνα 45 *- Η απεικόνιση των δεδομένων που συλλέχθηκαν για τον χρήστη SR.*

Περίπτωση 3 εικονικού χρήστη (YR)



Εικόνα 46 *- Η απεικόνιση των δεδομένων που συλλέχθηκαν για τον εικονικό χρήστη YR.*

Βιβλιογραφία

Amichai-Hamburger, Y., & Vinitzky, G. (2010). Social network use and personality. Computers in human behavior, 26(6), 1289-1295.

Association, A. H. I. M. (2001). Practice brief. Patient photography, videotaping, and other imaging (updated). Journal of AHIMA/American Health Information Management Association, 72(6), 64M.

Backstrom, L., Sun, E., & Marlow, C. (2010). Find me if you can: improving geographical prediction with social and spatial proximity. Paper presented at the Proceedings of the 19th international conference on World wide web.

Barr, K. E. (2007). ASIC Design in the Silicon Sandbox.

Barnes, N. G., & Lescault, A. M. (2012). The 2011 Inc. 500 social media update: Blogging declines as newer tools rule. Center for Marketing Research, University of Massachusetts, Dartmouth. Retrieved Feburary, 4.

Bean, H. (2007). The DNI's open source center: An organizational communication perspective. International Journal of Intelligence and Counterintelligence, 20(2), 240-257.

Best, C. (2011). Challenges in open source intelligence. Paper presented at the Intelligence and Security Informatics Conference (EISIC), 2011 European.

Bitcoin. (2014) Available from https://bitcoin.org/el/faq#what-is-bitcoin . [Accessed 25 September 2015]

Bradke, A. J. (2009). The Ethics of Medical Brigades in Honduras: Who are we helping? , University of Pittsburgh.

Brandtzæg, P. B., Lüders, M., & Skjetne, J. H. (2010). Too many Facebook “friends”? Content sharing and sociability versus the need for privacy in social network sites. Intl. Journal of Human–Computer Interaction, 26(11-12), 1006-1030.

Chandra, S., Khan, L., & Muhaya, F. B. (2011). Estimating twitter user location using social interactions--a content based approach. Paper presented at the Privacy, Security, Risk and Trust (PASSAT) and 2011 IEEE Third Inernational Conference on Social Computing (SocialCom), 2011 IEEE Third International Conference on.

Cohen, J. (1992). Statistical power analysis. Current directions in psychological science, 98-101.

Compton, R., Jurgens, D., & Allen, D. (2014). Geotagging one hundred million twitter accounts with total variation minimization. Paper presented at the Big Data (Big Data), 2014 IEEE International Conference on.

Compton, R., Lee, C.-K., Lu, T.-C., de Silva, L., & Macy, M. (2013). Detecting future social unrest in unprocessed twitter data: "emerging phenomena and big data". Paper presented at the Intelligence and Security Informatics (ISI), 2013 IEEE International Conference on.

Coughlan, T., & Perryman, L.-A. (2015). A Murky Business: Navigating the Ethics of Educational Research in Facebook Groups. Available from http://www.eurodl.org/materials/special/2015/Coughlan_Perryman.htm . [Accessed 30 September 2015]

Davolt Sr, D. (1999). The use of digital photography to support a medical mission to Honduras. The Journal of biocommunication, 27(2), 22-24.

Dhavase, N., & Bagade, A. (2014). Location Identification for Crime & Disaster Events by Geoparsing Twitter. 2014 International Conference for Convergence of Technology (I2CT), 22-24.

Du, H. S., & Wagner, C. (2006). Weblog success: Exploring the role of technology. International Journal of Human-Computer Studies, 64(9), 789-798.

Ellison, N. B., & Boyd, D. (2007). Social network sites: Definition, history, and scholarship. Journal of Computer-Mediated Communication, 13(1), 210-230.

European Derective. (2009). Available from http://ec.europa.eu/growth/sectors/electrical-engineering/rtte-directive . [Accessed 30 September 2015]

Eysenbach, G. (2008). Medicine 2.0: social networking, collaboration, participation, apomediation, and openness. Journal of medical Internet research, 10(3).

Eysenbach, G., & Till, J. E. (2001). Ethical issues in qualitative research on internet communities. Bmj, 323(7321), 1103-1105.

Facebook.com. (2015). Available from http://www.facebook.com . [Accessed 1 October 2015]

Fusek, J. (2015). Interoperable Process Design Kit a jeho automatizované generování.

Gakh, M. (2005). Argentina's protection of personal data: initiation and response. ISJLP, 2, 781.

Genymotion. (2014). Available from https://www.genymotion.com/#!/download [Accessed 1 September 2015]

Geopy.(2014) Available from https://github.com/geopy/geopy [Accessed 25 August 2015]

Geojonson. (2015). Available from http://geojson.org/geopy [Accessed 25 August 2015]

GoogleMaps. (2014). Available from https://developers.google.com/maps/documentation/javascript/ [Accessed 28 August 2015]

Geocoding. (2014). Available from https://developers.google.com/maps/documentation/geocoding/intro [Accessed 28 August 2015]

Google Geocoding API. (2014). Available from https://developers.google.com/maps/documentation/geocoding/intro [Accessed 28 August 2015]

Google Maps API. (2014). Available from https://developers.google.com/maps/pricing-and-plans/ [Accessed 28 August 2015]

Google Infowindows. (2014). Available from https://developers.google.com/maps/documentation/javascript/infowindows [Accessed 28 August 2015]

Google Polylines. (2014). Available from https://developers.google.com/maps/documentation/javascript/examples/polyline-simple [Accessed 28 August 2015]

Google Symbols. (2014). Available from https://developers.google.com/maps/documentation/javascript/symbols [Accessed 28 August 2015]

Gakh, M. (2005). Argentina's protection of personal data: initiation and response. ISJLP, 2, 781.

Hallaraker, O., & Vigna, G. (2005, June). Detecting malicious javascript code in mozilla. In Engineering of Complex Computer Systems, 2005. ICECCS 2005. Proceedings. 10th IEEE International Conference on (pp. 85-94). IEEE. Hanson, L., Harms, S., & Plamondon, K. (2011). Undergraduate international medical electives: some ethical and pedagogical considerations. Journal of Studies in International Education, 15(2), 171-185.

Hanson, L., Harms, S., & Plamondon, K. (2011). Undergraduate international medical electives: some ethical and pedagogical considerations. Journal of Studies in International Education, 15(2), 171-185.

Hansen, M., & Tschofenig, H. (2011). Terminology for Talking about Privacy by Data Minimization: Anonymity, Unlinkability, Undetectability, Unobservability, Pseudonymity, and Identity Management. draft-hansen-privacy-terminology-02 (work in progress).

Hulnick, A. S. (2010). The Dilemma of Open Sources Intelligence: Is OSINT Really Intelligence? Chapter, 14, 229.

Ignat, C., Steinberger, R., & Pouliquen, B. (2005). Navigating multilingual news collections using automatically extracted information. CIT. Journal of computing and information technology, 13(4), 257-264.

Java, A., Song, X., Finin, T., & Tseng, B. (2007). Why we twitter: understanding microblogging usage and communities. Paper presented at the Proceedings of the 9th WebKDD and 1st SNA-KDD 2007 workshop on Web mining and social network analysis.

Joosten, T. (2012). Social media for educators: Strategies and best practices: John Wiley & Sons.

Json. (2015), About. Available from http://www.json.org/ , [Accessed 28 August 2015]

Jurgens, D. (2013). That's What Friends Are For: Inferring Location in Online Social Media Platforms Based on Social Relationships. ICWSM, 13, 273-282.

Instagram. (2015), Blog. Available from http://blog.instagram.com/post/8756150468/a-real-time-api-for-next-generation-apps , [Accessed 1 October 2015]

Instagram. (2015), About. Available from http://instagram.com/about/us/ , [Accessed 1 October 2015]

Instagram. (2015), Developer. http://instagram.com/developer/ , [Accessed 1 October 2015]

Internet Engineering Task Force. (2014) The JavaScript Object Notation (JSON) Available from http://tools.ietf.org/html/rfc7159.html [Accessed 28 August 2015]

Kaplan, A. M., & Haenlein, M. (2010). Users of the world, unite! The challenges and opportunities of Social Media. Business horizons, 53(1), 59-68.

Lahneman, W. J. (2010). The need for a new intelligence paradigm. International Journal of Intelligence and Counterintelligence, 23(2), 201-225.

Lenhart, A., Purcell, K., Smith, A., & Zickuhr, K. (2010). Social Media & Mobile Internet Use among Teens and Young Adults. Millennials. Pew Internet & American Life Project.

Ma, C. Y., Yau, D. K., Yip, N. K., & Rao, N. S. (2013). Privacy vulnerability of published anonymous mobility traces. Networking, IEEE/ACM Transactions on, 21(3), 720-733.

Mandel, B., Culotta, A., Boulahanis, J., Stark, D., Lewis, B., & Rodrigue, J. (2012). A demographic analysis of online sentiment during hurricane irene. Paper presented at the Proceedings of the Second Workshop on Language in Social Media.

Masoumzadeh, A., & Joshi, J. (2012). Preserving structural properties in edge-perturbing anonymization techniques for social networks. Dependable and Secure Computing, IEEE Transactions on, 9(6), 877-889.

Memon, M. H., Khan, A., Li, J.-P., Shaikh, R. A., Memon, I., & Deep, S. (2014). Content based image retrieval based on geo-location driven image tagging on the social web. Paper

presented at the Wavelet Active Media Technology and Information Processing (ICCWAMTIP), 2014 11th International Computer Conference on.

Mocanu, D., Baronchelli, A., Perra, N., Gonçalves, B., Zhang, Q., & Vespignani, A. (2013). The twitter of babel: Mapping world languages through microblogging platforms. PloS one, 8(4), e61981.

Mok, D., Wellman, B., & Carrasco, J. (2010). Does distance matter in the age of the Internet? Urban Studies, 47(13), 2747-2783.

Ong, S. P., Cholia, S., Jain, A., Brafman, M., Gunter, D., Ceder, G., & Persson, K. A. (2015). The Materials Application Programming Interface (API): A simple, flexible and efficient API for materials data based on REpresentational State Transfer (REST) principles. Computational Materials Science, 97, 209-215.

Paar, C., & Pelzl, J. (2009). Understanding cryptography: a textbook for students and practitioners. Springer Science & Business Media.

Papacharissi, Z. (2009). The virtual geographies of social networks: a comparative analysis of Facebook, LinkedIn and ASmallWorld. New media & society, 11(1-2), 199-220.

Parsell, M. (2008). Pernicious virtual communities: Identity, polarisation and the Web 2.0. Ethics and Information Technology, 10(1), 41-56.

Paul, M. J., & Dredze, M. (2011). You are what you Tweet: Analyzing Twitter for public health. Paper presented at the ICWSM.

Peng, W., Li, F., Zou, X., & Wu, J. (2014). A two-stage deanonymization attack against anonymized social networks. Computers, IEEE Transactions on, 63(2), 290-303.

Python Developer. (2015). Available from https://docs.python.org/devguide/ [Accessed 28 August 2015]

Python Library. (2015). Available https://docs.python.org/2/library/csv.html [Accessed 28 August 2015]

Python Codecs, (2014). Available https://docs.python.org/2/library/codecs.html [Accessed 28 August 2015]

Ramim, M., & Levy, Y. (2006). Securing e-learning systems: A case of insider cyber attacks and novice IT management in a small university. Idea Group Publishing.

Quan-Haase, A., & Young, A. L. (2010). Uses and gratifications of social media: A comparison of Facebook and instant messaging. Bulletin of Science, Technology & Society, 30(5), 350-361.

Rakesh, V., Reddy, C. K., Singh, D., & Ramachandran, M. (2013). Location-specific tweet detection and topic summarization in twitter. Paper presented at the Advances in Social Networks Analysis and Mining (ASONAM), 2013 IEEE/ACM International Conference on.

Ramim, M., & Levy, Y. (2006). Securing e-learning systems: A case of insider cyber attacks and novice IT management in a small university. Idea Group Publishing.

Rodrigues, R. J. (2000). Ethical and legal issues in interactive health communications: a call for international cooperation. Journal of Medical Internet Research, 2(1).

Takhteyev, Y., Gruzd, A., & Wellman, B. (2012). Geography of Twitter networks. Social networks, 34(1), 73-81.

Thompson, L. A., Black, E., Duff, W. P., Black, N. P., Saliba, H., & Dawson, K. (2011). Protected health information on social networking sites: ethical and legal considerations. Journal of medical Internet research, 13(1).

Tumasjan, A., Sprenger, T. O., Sandner, P. G., & Welpe, I. M. (2010). Predicting Elections with Twitter: What 140 Characters Reveal about Political Sentiment. ICWSM, 10, 178-185.

Tweepy API (2015). Available from http://tweepy.readthedocs.org/en/v3.2.0/getting_started.html#api [Accessed 28 August 2015]

Twitter. (2015), emarketer. Available from https://twitter.com/emarketer , [Accessed 2 October 2015]

Twitter. (2015), company Available from https://twitter.com/ccompany , [Accessed 4 October 2015]

Twitter. (2014), RestAPI Available from https://dev.twitter.com/rest/public https://dev.twitter.com/rest , [Accessed 9 October 2015]

Twitter. (2014), GeoJson Available from http://www.geojson.org/ , [Accessed 4 October 2015]

Twitter. (2014), StreamingAPI Available from https://dev.twitter.com/streaming/overview , [Accessed 3 October 2015]

Twitter. (2014), OAUTH Available from https://dev.twitter.com/oauth , [Accessed 4 October 2015]

Twitter. (2014), Timeline Available from https://dev.twitter.com/rest/reference/get/statuses/user_timeline , [Accessed 4 October 2015]

Twitter IDs. (2014), Snowflake, Available from https://dev.twitter.com/overview/api/twitter-ids-json-and-snowflake [Accessed 4 October 2015]

Valkanas, G., & Gunopulos, D. (2012). Location extraction from social networks with commodity software and online data. Paper presented at the Data Mining Workshops (ICDMW), 2012 IEEE 12th International Conference on.

Washington Post. (2013), Available from https://www.washingtonpost.com/business/technology/twitter-turns-7-users-send-over-400-million-tweets-per-day/2013/03/21/2925ef60-9222-11e2-bdea-e32ad90da239_story.html , [Accessed 1 October 2015]

Weitzman, E. R., Kaci, L., & Mandl, K. D. (2010). Sharing medical data for health research: the early personal health record experience. Journal of medical Internet research, 12(2).

Xu, L., Jiang, C., Wang, J., Yuan, J., & Ren, Y. (2014). Information Security in Big Data: Privacy and Data Mining. Access, IEEE, 2, 1149-1176.

Yamaguchi, Y., Amagasa, T., & Kitagawa, H. (2013). Landmark-based user location inference in social media. Paper presented at the Proceedings of the first ACM conference on Online social networks.

Yuan, M., Chen, L., Yu, P. S., & Yu, T. (2013). Protecting sensitive labels in social network data anonymization. Knowledge and Data Engineering, IEEE Transactions on, 25(3), 633-647..

Zhang, L., Li, X., Liu, K., Jung, T., & Liu, Y. (2014). Message in a Sealed Bottle: Privacy Preserving Friending in Mobile Social Networks.

Αρχή Προστασίας Δεδομένων Προσωπικού Χαρακτήρα, (2009). Available from http://www.dpa.gr/portal/page?_pageid=33,6948&_dad=portal&_schema=PORTAL [Accessed 7 October 2015]

EETT (2009). (2009). Available from http://www.eett.gr/opencms/opencms/EETT/Electronic_Communications/Telecoms/SupplierSubscriberIdentification/ , [Accessed 7 October 2015]

Τόμουζου, Λ. (2013). Τεχνολογίες ταυτοποίησης ατόμου με σεβασμό της ιδιωτικότητας του: Θεωρία και εφαρμογές.