Τεχνολογικό Εκπαιδευτικό Ίδρυμα Κρήτης

Σχολή Τεχνολογικών Εφαρμογών

Τμήμα Μηχανικών Πληροφορικής

Πτυχιακή Εργασία

Τίτλος: Μελέτη, Χρήση και Εφαρμογή των CAPTCHA

ΓΕΩΡΓΙΟΣ ΣΠΑΡΗΣ

Επιβλέπων Καθηγητής:ΧΑΡΗΣ ΜΑΝΙΦΑΒΑΣ

Επιτροπή Αξιολόγησης:

Περιεχόμενα

ΠΙΝΑΚΑΣ ΕΙΚΟΝΩΝ

ΠΕΡΙΛΗΨΗ

Τα CATCHAs είναι μια δημοφιλής μέθοδος για τη ανακοπή αυτοματοποιημένων επιθέσεων και να μειωθεί το spam σε ιστοσελίδες. Η ιδέα είναι να χρησιμοποιηθεί από τα άτομα που αναπτύσσουν ιστοσελίδες, ορισμένα παζλ που μόνο ένα ανθρώπινο ον μπορεί να λύσει, αλλά όχι αυτοματοποιημένα προγράμματα.Ιδιαίτερα τα CAPTCHAs κειμένου, που ζητούν από το χρήστη να αποκρυπτογραφήσει παραμορφωμένο κείμενο σε μια εικόνα, είναι ευρέως διαδεδομένα στις ιστοσελίδες και οι χρήστες πλέον είναι εξοικειωμένοι με αυτό το είδος της CAPTCHA. Η reCAPTCHA είναι ένας από τους μεγαλύτερος πάροχος του κειμένου με βάση την τεχνολογία CAPTCHA και μάλιστα θεωρείται ως ένας από τους πιο ασφαλείς. Οι τρέχουσα οπτικές μεθόδους αναγνώρισης χαρακτήρων, ιδιαίτερα οι παραδοσιακές μεθόδοι κατάτμησης, δεν μπορούν να κατασταθούν αποτελεσματικές στις παραμορφωμένες λέξεις από το reCAPTCHA. Σε αυτή τη μελέτη, η ασφάλεια των CAPTCHA τίθεται υπό αμφισβήτηση αφού αφενός θα αναλυθούν τα πλεονεκτήματα και οι λόγοι που οδήγησαν στην ανάπτυξη του μηχανισμού ενώ ταυτόχρονα θα αξιολογηθεί η αποτελεσματικότητά του μέσα από την εφαρμογή σε ιστοσελίδα που ανέπτυξε ο ίδιος ο συντάκτης.

Λέξεις κλειδιά: CAPTCHA, Ασφάλεια, Αποκωδικοποίηση CAPTCHA, Spamming, Web

Abstract

CAPTCHAs are a popular method for stopping automated attacks and to reduce spam on websites. The idea is to use puzzle, that only a human being can solve, but not automated programs. Particularly text CAPTCHAs, that ask a ser to decipher distorted text in an image, are widely deployed on websites and users are accustomed to this type of CATCHA. reCAPTCHA is one of the biggest provider of text based CAPTCHA technology and is said to be one of the most secure. Current optical character recognition methods, especially traditional segmentation methods, work unreliably on distorted words from reCAPTCHA. In this study, the safety of CAPTCHA questioned since both author will analyze the advantages and the reasons that led to the development of the mechanism while evaluating effectiveness through the implementation of a website developed by the author himself.

Keywords: CAPTCHA, Internet security, Decoding CAPTCHA, Spamming, Web

Ευχαριστίες

Καταρχήν θα ήθελα να ευχαριστήσω τους καθηγητές μου στο Τμήμα Εφαρμοσμένης Πληροφορικής και Πολυμέσων και ιδιαίτερα τον καθηγητή μου, Χάρη Μανιφάβα, που μου ανέθεσε το συγκεκριμένο θέμα για την πτυχιακή μου εργασία καθώς και για το χρόνο που μου αφιέρωσε προκειμένου να την ολοκληρώσω. Οι ουσιαστικές παρατηρήσεις του με οδήγησαν στη βαθύτερη κατανόηση του συγκεκριμένου θέματος και συνέβαλαν στη δημιουργία του παρόντος πονήματος.

Τέλος θα ήθελα να ευχαριστήσω την οικογένειά μου για την αμέριστη συμπαράσταση που μου προσέφερε σε όλη την πορεία μου μέχρι σήμερα χωρίς να μου στερήσει οποιαδήποτε ψυχολογική και υλική αρωγή.

ΕΙΣΑΓΩΓΗ - ΣΚΟΠΟΣ ΤΗΣ ΕΡΓΑΣΙΑΣ

Κάθε χρήστης του διαδικτύου, κατά τη διάρκεια της περιήγησης του, έχει αναγκαστεί να σταματήσει και να εκτελέσει μία μικρή περίεργη διεργασία που περιλαμβάνει την αντιγραφή κάποιων αριθμών, γραμμάτων και συμβόλων που είναι μάλλον παραμορφωμένα σε ένα κενό πεδίο. Γενικά η διαδικασία αυτή μοιάζει να είναι ανούσια για τον χρήστη, αλλά σίγουρα δεν είναι για τους ανθρώπους που βρίσκονται πίσω από την εφαρμογή. Αυτή η σειρά γραμμάτων ακούει στο όνομα CAPTCHA και στην ουσία αποτελεί μία απλή δοκιμασία για να αποδειχτεί ότι ο επισκέπτης της σελίδας δεν είναι ένας υπολογιστής, αλλά ένας άνθρωπος. Κινητήρια δύναμη πίσω από την ανάπτυξη του μηχανισμού υπήρξε η εμφάνιση μιας αυτοματοποιημένης διαδικασίας που δημιουργούσε spam mails δρώντας μέσα στο Yahoo.

Η επιτυχία του συγκεκριμένου μηχανισμού αναγνώρισης ήταν άμεση. Σήμερα εκατομμύρια άνθρωποι παγκοσμίως εξαναγκάζονται να ακολουθήσουν τη διαδικασία προσπαθώντας να διαχωρίσουν τη θέση τους ανάμεσα στον αόρατο πόλεμο που διεξάγεται μεταξύ των spammers και των προγραμματιστών. Η δημιουργία του μηχανισμού των CAPTCHA δεν αποτέλεσε τον τερματισμό των επιχειρήσεων των spammers, αλλά μάλλον την απαρχή για ενός νέου πεδίου μάχης. Οι υπολογιστές βελτιώνονται όλο και περισσότερο, γεγονός που αναγκάζει τους προγραμματιστές να δημιουργήσουν πιο εξελιγμένους μηχανισμούς. Η διαδικασία είναι αέναη και συνεχώς εξελισσόμενη γεγονός που δημιουργεί μία εντεινόμενη αντιπαλότητα στον τομέα της ασφάλειας.

Η παρούσα εργασία προσπαθεί να κινηθεί μέσα σε τρία επίπεδα προκειμένου να καταστήσει κατανοητή τη λογική και την τεχνολογία του μηχανισμού. Σε πρώτο επίπεδο θα επιχειρηθεί να αναλυθεί η αναγκαιότητα του μηχανισμού CAPTCHA, τα κίνητρα που ώθησαν στην ανάπτυξη του και τις υπάρχουσες λύσεις που έχει δημιουργήσει η τεχνολογία. Ποια είναι τα προβλήματα στα οποία έδωσε απάντηση το CAPTCHA; Μέσα από το συγκεκριμένο τμήμα της εργασίας θα αναλυθούν βασικά ιστορικά στοιχεία και ταυτόχρονα θα αναδεχθούν τα πλεονεκτήματα και τα μειονεκτήματα του μηχανισμού.

Σε δεύτερο επίπεδο θα δημιουργηθεί μία ιστοδελίδα με χρηστικό χαρακτήρα η ασφάλεια της οποία θα βασίζεται στην ανάπτυξη ενός CAPTCHA συστήματος. Στόχος είναι η δημιουργία του αντικειμένου που χρήζει ασφάλειας και ταυτόχρονα να δημιουργηθεί η υποδομή για την εξασφάλισή της. Έχοντας θέσει τον μηχανισμό ασφαλείας θα αναπτυχθεί ένας μηχανισμός ο οποίος να μπορεί να αποκωδικοποιεί των μηχανισμό CAPTCHA.

Η αναγκαιότητα μιας τέτοιας μελέτης είναι δεδομένη καθώς θα επιχειρηθεί να καλυφθεί το ζήτημα τόσο σε θεωρητικό όσο και σε πρακτικό επίπεδο. Η βιβλιογραφία για το συγκεκριμένο θέμα στον ελληνικό χώρο είναι σίγουρα ελλειμματική ενώ ταυτόχρονα η υπάρχουσα στην αγγλική είναι πλούσια αλλά αποσπασματική, καλύπτοντας κάθε φορά μόνο ένα μέρος του ζητήματος. Η απουσία μίας εμπεριστατωμένης μελέτης συνδυάζοντας την πρακτική εφαρμογή με το θεωρητικό υπόβαθρο ουσιαστικά απουσίαζε από την ελληνική βιβλιογραφία.

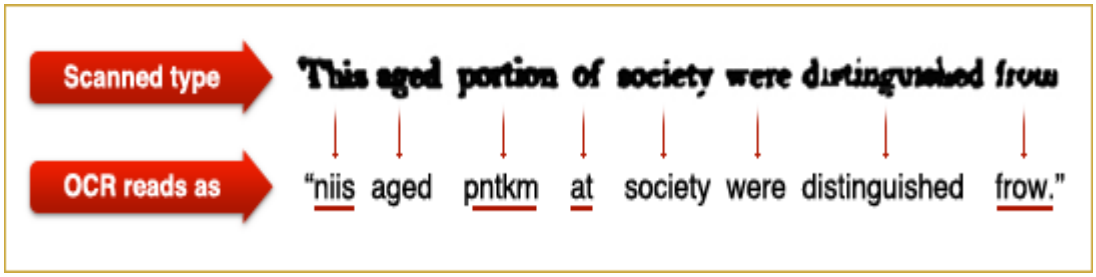
Στόχος του συντάκτη της εργασίας είναι να αναλυθούν επιστημονικά άρθρα από έγκριτες βάσεις δεδομένων, αφού η έντυπη βιβλιογραφία αφενός είναι φτωχή αφετέρου η ηλεκτρονική βιβλιογραφία είναι πιο επίκαιρη. Η χρήση επίκαιρων πηγών θα οδηγήσει στη δημιουργία ενός επίκαιρου πονήματος.

– ΕΙΣΑΓΩΓΗ

1.1. ΓΕΝΙΚΑ

Το CAPTCHA είναι μία αυτόματη δοκιμασία η οποία εφαρμόζεται στο διαδίκτυο και χρησιμοποιείται για να κατασταθεί ικανός ο διαχωρισμός μεταξύ ανθρώπινων ή μηχανικών επισκεπτών. Ο όρος έχει δημιουργηθεί από τα αρχικά "Completely Automated Public Turing test to tell Computers and Humans Apart" και δημιουργήθηκε από τους Ahn και συν. το 2000 (Hopper και συν., 2000). Ένα CAPTCHA θα πρέπει να είναι εύκολο στο να επιλυθεί από έναν άνθρωπο και δύσκολο για ένα λογισμικό. Ένα λογισμικό κρύβεται πίσω από τη δοκιμασία το οποίο μπορεί να διακρίνει τη φύση του επισκέπτη (Coates et al., 2001). Η πιο συνηθισμένη εφαρμογή ενός CAPTCHA είναι η αποτροπή του spamming (Hopper και συν., 2000). Διαδικτυακό Bots και κακόβουλοι χρήστες μπορούν να δημιουργήσουν χιλιάδες ψεύτικα σχόλια και να κατακλύσουν σελίδες με spam μηνύματα. Ταυτόχρονα στόχος των κακόβουλων εφαρμογών γίνονται και διαδικτυακές δημοσκοπήσεις με σκοπό να χειραγωγούνται τα αποτελέσματα. Οι διαχειριστές των ιστοσελίδων μπορούν να χρησιμοποιήσουν τέτοιους μηχανισμούς προκειμένου να εξακριβώσουν την ταυτότητα των επισκεπτών και να επιτρέψουν ή όχι την παροχή συγκεκριμένων υπηρεσιών.

Η λογική της δημιουργίας των CAPTCHA στηρίζεται στην παραδοχή ότι η πληροφορική δεν έχει φτάσει ακόμα στο επίπεδο που απαιτείται για να κάνει το συγκεκριμένο διαχωρισμό. Οι μηχανισμοί των CAPTCHA βασίζονται στην παραμόρφωση εικόνων χαρακτήρων και λέξεων προκειμένου να αυξήσουν το βαθμό δυσκολίας του εγχειρήματος. Ακόμα και η τεχνολογία της οπτικής αναγνώρισης χαρακτήρων (Optical Character Recognition - OCR) δεν έχει φτάσει ακόμα στο κατάλληλο επίπεδο για να ξεπεράσει το βαθμό δυσκολίας. Για την ώρα οι άνθρωποι θεωρούνται πιο ισχυροί από την τεχνολογία στο συγκεκριμένο ζήτημα και μάλιστα όταν υπάρχει αυξημένος βαθμός παραμόρφωσης του κειμένου που πρέπει να αντιγραφεί. Chellapilla και συν., (2005) υποστηρίζουν ότι χαρακτήρες μεμονωμένοι μπορούν να αναγνωριστούν από τα υπολογιστικά συστήματα, αλλά ένα σύνολο είναι ακόμα δύσκολο.



Εικόνα 1-1 Αναγνώριση CAPTCHA από OCR

Η τεχνολογία OCR μπορεί να απομονώσει χαρακτήρες αλλά σίγουρα δεν μπορεί να έχει ανάλογα αποτελέσματα και σε συνδυασμούς χαρακτήρων (Chellapilla και Simard, 204).

Η όλη προσέγγιση του CAPTCHA δεν σχετίζεται αποκλειστικά με τη δυνατότητα της μηχανής να αναγνωρίζει χαρακτήρες και λέξεις. Ο ανθρώπινος παράγοντας είναι εξίσου σημαντικός, αφού και το άτομο παίζει ρόλο στην όλη διαδικασία. Η αναγνώριση λέξεων από τον άνθρωπο απετέλεσε πεδίο μελέτης της ψυχολογίας κατά τον περασμένο αιώνα (Grainger και Jacobs, 1996). Δύο ήταν οι βασικές θεωρίες που αναπτύχθηκαν σχετικά με την αντιληπτική ικανότητα του ανθρώπου:

- Η ολιστική που θεωρεί ότι το άτομο μπορεί να αναγνωρίζει ολόκληρες λέξεις με τη μία
- Η ιεραρχική που στηρίζεται στη λογική της τμηματοποίησης της λέξεις σε χαρακτήρες ή σε κομμάτια.

Ο Frost και συν. (1998) θεωρούν ότι και οι δύο θεωρίες έχουν λογική βάση και μάλιστα ότι το άτομο λειτουργεί συνδυαστικά. Τα πειράματα των τελευταίων βασίστηκαν σε άτομα που πάσχουν από δυσλεξία. Οι Lavrenko και συν. (2004) θωρεί ότι η ολιστική θεωρία είναι πιο ισχυρή και ότι η αναγνώριση γίνεται αυτόματα χωρίς καμία τμηματοποίηση.

Το reCAPTCHA είναι ένα από τα πιο συνηθισμένα CAPTCHA που χρησιμοποιούν οι διαχειριστές ιστοσελίδων. Μέχρι σήμερα έχει βασιστεί επί το πλείστον στην αναγνώριση λέξεων. Οι λέξεις που χρησιμοποιούνται προέρχονται από βιβλία και εφημερίδες που έχουν σαρωθεί και οι περισσότερες είναι σπάνιες προκειμένου να μην έχουν υποβληθεί ως test set σε κάποια μέθοδο machine learning. Ταυτόχρονα η εμφάνισή τους δεν είναι η συνηθισμένη, αφού χρησιμοποιούνται παράξενα και ποικίλα φόντα ακόμα και στην ίδια τη λέξη προκειμένου να μην μπορούν να κατασταθούν αναγνωρίσιμα από μεθόδους OCR. Μία άλλη μέθοδος που χρησιμοποιείται για την αύξηση της δυσκολίας είναι η δημιουργία πολύ μικρών κενών ανάμεσα στις λέξεις. Η διαδικασία αναγνώρισης από τους μηχανισμού λογισμικού καθίσταται ακόμα πιο δύσκολη. Η παρακάτω εικόνα παρουσιάζει ένα χαρακτηριστικό παράδειγμα CAPTCHA. Όπως παρατηρείτε η γραμματοσειρά δεν είναι συνηθισμένη. Οι μαύρες κηλίδες που περιβάλλουν τις λέξεις αποδεικνύουν ότι έχει εφαρμοστεί ένας συγκεκριμένος βαθμός δυσκολίας.

Εικόνα 1-2: Παράδειγμα CAPTCHA

Η ιστοσελίδα του reCAPTCHA αναφέρει ότι καθημερινά δημιουργεί 30 εκατομμύρια προκλήσεις. Η ενσωμάτωση ενός reCAPTCHA σε μία ιστοσελίδα είναι δωρεάν σε διάφορες γλώσσες προγραμματισμού καθιστώντας την υιοθέτηση της συγκεκριμένης τεχνολογίας πολύ εύκολη. Όπως αναφέρει η ιστοσελίδα reCAPTCHA αυτή τη στιγμή περισσότερα από 100.000 ιστοσελίδες χρησιμοποιούν τη συγκεκριμένη τεχνολογία. Ανάμεσα σε αυτές είναι τα Facebook, Twitter και StumbleUpon, οι οποίες έσπευσαν να υιοθετήσουν ένα CAPTCHA από το 2007. Δεν είναι λίγα τα ακαδημαϊκά ιδρύματα που έχουν επιλέξει ένα CAPTCHA για την ασφάλειά τους. Ταυτόχρονα η αναζήτηση μεθόδων που να μπορούν να αγνοούν τη τεχνολογία αυτή αποτελεί πάντα ένα από τα βασικά ερευνητικά ζητούμενα των τμημάτων πληροφορικής. Στα πλαίσια αυτού του ανοιχτού επιστημονικού πεδίου η παρούσα πτυχιακή εργασία θα δραστηριοποιηθεί στην ανάπτυξη ενός μηχανισμού που θα μπορεί να υλοποιεί το συγκεκριμένο στόχο.

1.2. ΤΟ ΠΡΟΒΛΗΜΑ

Οι ιστοσελίδες που μπορούν να προσελκύσουν διάφορες μορφές αυτοματισμού έχουν λάβει μέτρα προστασίας προκειμένου να διασφαλίσουν την ανθρώπινη υπόσταση των επισκεπτών τους, όπως αναφέρθηκε και παραπάνω. Μπορεί το όνομα CAPTCHA να δόθηκε από τους ανθρώπους του Πανεπιστημίου Carnegie Mellon στον μηχανισμό, αλλά στη συνέχεια το όνομα έτυχε ευρείας αποδοχής, αφού όλα τα projects που βασίζονταν στην ίδια λογική ονομάστηκαν έτσι. Στην ουσία ο ορισμός αγκαλιάζει μία συγκεκριμένη μέθοδο και όχι ένα μεμονωμένο project. Μπορεί να έχει χαρακτηριστεί ως επιτυχημένο, αλλά στην ουσία δεν μπορεί να εξυπηρετήσει άτομα που έχουν προβλήματα όρασης ή είναι δυσλεκτικοί.

Σίγουρα όπως όλα τα συστήματα ασφαλείας δεν απολαμβάνουν την απολυτότητα. Σε όλα υπάρχει ένα μειονέκτημα. Για παράδειγμα η συγκεκριμένη τεχνολογία μπορεί να υιοθετηθεί από spammers πληρώνοντας ανθρώπους να αναγνωρίζουν λέξεις που στη συνέχεια θα δίνονται σε ένα λογισμικό που θα τις χρησιμοποίει ως παράδειγμα για να σπάει συστήματα CAPTCHA. Ορισμένες ιστοσελίδες μεγάλων οργανισμών ή τραπεζών εργάζονται προς την τελειοποίηση τους, όπως για παράδειγμα η ING Direct που ονομάζει το μηχανισμό της ως

«PIN Guard» που προσπαθεί να εξαλείψει τα προβλήματα που αντιμετωπίζουν χρήστες με μειωμένη όραση. Μπορεί οι μεγάλες ιστοσελίδες να πληρώνουν αδρά για να βελτιστοποιήσουν τέτοια συστήματα, αλλά και οι απλοί bloggers χρησιμοποιούν το CAPTCHA, αφού όπως ισχυρίζονται μειώνουν στο ελάχιστο το spamming.

Όπως αναφέρει η W3 (http://www.w3.org) το CAPTCHA δε μπορεί να θεωρηθεί πανάκεια για το πρόβλημα του spamming. Την ίδια στιγμή που μία ομάδα δημιουργούσε το CAPTCHA στο Πανεπιστήμιο του Carnegie Mellon, ταυτόχρονα μία άλλη εργαζόταν για να εντοπίσει προβλήματα και να το σπάσει. Μάλιστα η πρώτη επιτυχημένη προσπάθεια για να σπάσει ο μηχανισμός έγινε από φοιτητή του ίδιου ιδρύματος. Είναι κατανοητό λοιπόν ότι το CAPTCHA μπορεί να έχει υψηλό βαθμό επιτυχίας, αλλά δεν έχει εξασφαλίσει την απόλυτη.

1.3. ΙΣΤΟΡΙΚΑ ΣΤΟΙΧΕΙΑ

Η ενότητα αυτή θα αναφερθεί επιγραμματικά με εργασίες που έχουν γίνει πάνω σε αυτό το αντικείμενο προκειμένου να έχει ο αναγνώστης ξεκάθαρη εικόνα σχετικά με την πορεία ανάπτυξης της συγκεκριμένης τεχνολογίας. Το πρώτο πρακτικά εφαρμοζόμενο CAPTCHA σύστημα βασιζόταν σε κείμενο και αναπτύχθηκε το 1998 από την Compaq Computer Corp. (Lillibridge, 2005). Το 2000 ο Ahn και οι συνεργάτες του εισήγαν την έννοια του CAPTCHA όπως τη γνωρίζουμε εμείς σήμερα και το σχέδιο αυτό ολοκληρώθηκε το 2003 με την τεκμηριωμένη περιγραφή του. Οι ίδιοι στη συνέχει εργάστηκαν για την εταιρεία Yahoo προκειμένου να δημιουργήσουν γι αυτή κάποιες μορφές CAPTCHA, όπως είναι αυτό του EZ-Gimpy και Gimpy. Το 1999 οι Nagy και συν. (1999) απέδειξαν ότι οι ικανότητες της οπτικής αναγνώρισης χαρακτήρων από λογισμικών δεν μπορούσαν να συγκριθούν με αυτές ενός επτάχρονου παιδιού.

Ο Coates και συν. (2001) βασίστηκε στα προαναφερθέντα αποτελέσματα και δημιούργησε μία περιγραφή ενός συστήματος CAPTCHA όπου οι χαρακτήρες δεν θα παρουσιάζονταν στη συνηθισμένη τους μορφή, αλλά σε κυματοειδή και μάλιστα σε συνδυασμό με χαμηλής ποιότητας αποτύπωση. Το 2003 οι Simard και συν. (2003) θεώρησε ότι με βάση τα πειράματά του το θέμα της οπτικής αναγνώρισης χαρακτήρων έχει λυθεί. Παρόλα αυτά το ζήτημα της τμηματοποίησης των χαρακτήρων είναι υπαρκτό και στόχος των τεχνολογιών CAPTCHA είναι να προσπαθήσουν να εστιάσουν στην αποτροπή της τμηματοποίησης για να αυξηθεί ο βαθμός δυσκολίας. Στην ουσία εισήγαγαν στο λεξιλόγιο της τεχνολογίας την έννοια του segmentation resistant συστήματος.

Οι Baird και συν. (2005) δημιούργησαν το ScatterType το 2005, που αποτελεί ένα σύστημα CAPTCHA το οποίο χρησιμοποιεί λέξεις που είναι τεχνητά δημιουργημένες προκειμένου να μπορούν να αντιστέκονται στην παραδοσιακή τμηματοποίηση που κάνουν τα OCR. Ο Ahn και οι συνεργάτες του εισάγουν το reCAPTCHA το 2008, ένα σύστημα CAPTCHA ο οποίο χρησιμοποιεί λέξεις από παλαιά βιβλία και βοηθάει στην ψηφιοποίησή τους. Σύμφωνα με τις πρώτες δοκιμές το νέο σύστημα μοιάζει να είναι ανθεκτικό σε επιθέσεις τμηματοποίησης λέξεων και γι αυτό το λόγο κέρδισε και την εμπιστοσύνη του κοινού πολύ άμεσα. Πλέον το ενδιαφέρον της επιστημονικής κοινότητας έχει στραφεί σε μεγάλο βαθμό στην τμηματοποίηση των χαρακτήρων που θεωρείται ως το βασικό όπλο για το σπάσιμο του συγκεκριμένου τύπου ασφαλείας.

Από τη στιγμή της εισαγωγής της νέας τεχνολογίας, τα CAPTCHAs κειμένου αποτελούν την πιο δημοφιλή μορφή και γι αυτό το λόγο η υφιστάμενη βιβλιογραφία στοχεύει όσο το δυνατόν περισσότερο σε αυτό το σημείο. Το 2003 έγινε προσπάθεια να περιγραφεί ένα αυτοματοποιημένο σύστημα το οποίο θα διασπούσε την ασφάλεια των EZ-Gimpy και Gimpy CAPTCHA (Mori και Malik, 2003). Το πρώτο διασπάστηκε με ποσοστό επιτυχίας που άγγιξε το 33%, ενώ το δεύτερο σχεδόν άγγιξε την τελειότητα με το 92%. Η λογική πάνω στην οποία βασίστηκαν τα μοντέλα διάσπασης ήταν η προσέγγιση του σχήματος των συμβόλων μέσα από μαθηματικά μοντέλα. Η ιδέα αναπτύχθηκε από πολλούς επιστήμονες και μάλιστα μετεξελίχθηκε. Οι Chellapilla και Simard (2004) έσπασαν πολλούς CAPTCHA μηχανισμούς χρησιμοποιώντας διάφορες λογικές. Ο αλγόριθμος αναγνώρισης των χαρακτήρων είχε πολύ καλά ποσοστά επιτυχίας. Η χρήση αλγορίθμων μηχανικής εκμάθησης ήταν το μυστικό της επιτυχίας. Ταυτόχρονα οι Chellapilla και Simard (2004) σημείωσαν ότι όσο μεγαλύτερη είναι η παραμόρφωση των χαρακτήρων τόσο μικρότερη είναι η πιθανότητα να σπάσει το CAPTCHA. Το 2005 ο Chellapilla και συν. (2005) αμφισβήτησαν το γεγονός ότι οι παραμορφωμένοι χαρακτήρες δύσκολα αναγνωρίζονται μέσα από ένα σύνολο δοκιμών. Αντίθετα ο αλγόριθμος που ανέπτυξαν μπορούσε να δώσει αποτελέσματα καλύτερα και από την ανθρώπινη αντίληψη. Στον πίνακα που ακολουθεί μπορούμε να παρακολουθήσουμε τα ποσοστά επιτυχίας ανάλογα με την παραμόρφωση.

Characters under typical distortions	Recognition rate of computers
	≈ 100%
	96+%
	100%
	98%
	≈ 100%
	95+%

Εικόνα 1-3 Ποσοστά επιτυχίας ανάλογα με το βαθμό παραμόρφωσης

Πολλές από τις πιο πρόσφατες προσεγγίσεις θεωρούν ως βάση την τμηματοποίησης της λέξης σε χαρακτήρες και στόχευση στην αναγνώριση του καθενός. Το 2009 ο Wilkins (2009) δημοσίευσε ένα σύνολο κανόνων που θα θωράκιζαν τα αναπτυσσόμενα συστήματα από πιθανές επιθέσεις.

1.4. ΣΥΝΟΨΗ ΚΕΦΑΛΑΊΟΥ

Είναι δεδομένο ότι η εποχή του διαδικτύου δεν επέφερε μόνο θετικά στοιχεία στη ζωή των χρηστών, αλλά δημιούργησε και αμέτρητους πονοκεφάλους στα άτομα που ασχολούνται με την ανάπτυξη εφαρμογών που θα προστατεύσουν την εγκυρότητα και τις συνδιαλλαγές που διενεργούνται στο νέο εργαλείο. Το CAPTCHA δημιουργήθηκε με σκοπό να αναγνωρίσει το ανθρώπινο χρήστη από το ανεπτυγμένο λογισμικό. Η αρχική μορφή του CAPTCHA θεωρείται πλέον παρωχημένη καθώς έχουν αναπτυχθεί τα αντίδοτα στους περιορισμούς που αυτοί έθεταν. Όσο πιο εξελιγμένοι μηχανισμοί αναπτύχθηκαν τόσο πιο πολύ έπρεπε να αναπτυχθεί και η λογική στην οποία έπρεπε να βασιστεί ο μηχανισμός αναγνώρισης. Στο κεφάλαιο που ακολουθεί θα μελετηθούν όλες οι εξελίξεις στον τομέα. των CAPTCHA. Στο παιχνίδι μπήκαν πλέον πέρα από το κείμενο και οι εικόνες. Το ανθρώπινο μάτι είναι σαφέστατα πιο έμπειρο από το μηχανικό και μπορεί να αναγνωρίζει ανωμαλίες και παραμορφώσεις. Το πεδίο ανάπτυξης των μηχανισμών ασφαλείας αποτελεί το αντίπαλο δέος των μηχανισμών παραβίασης γεγονός που θα μελετηθεί σε επόμενο κεφάλαιο.

ΠΗΓΕΣ ΚΕΦΑΛΑΙΟΥ

- Nicholas J., Ahn Hopper Luis Von Blum, Manuel, and Langford. John (2000) The official CAPTCHA site

- Coates, AL , Baird, HS and Fateman. RJ (2001) Pessimal print: a reverse Turing test. *Proceedings of the Sixth International Conference on Document Analysis and Recognition*

- Chellapilla, K. Larson, K.. Simard , P and Czerwinski M.. (2005) Computers beat humans at single character recognition in reading based human interaction proofs (HIPs). In *Proceedings of the Second Conference on Email and Anti-Spam*, pages 21–22. Citeseer

- Chellapilla K. Simard. P. (2004)Using machine learning to break visual human interaction proofs (HIPs). Advances in Neural Information Processing Systems, 17

- Chew M and Tygar JD. (2005) Image recognition CAPTCHAs. *Information Security*

- Chellapilla, K. Larson, K. Simard, P.Y. and Czerwinski M.. (2005) Building segmentation based human-friendly human interaction proofs (HIPs). *Human Interactive Proofs*, pages 1–26

- Lillibridge, MD. Abadi, M. Bharat, K. and Border A.. (2001) Method for selectively restricting access to computer systems. *US Patent 6,195,698*

- Nagy, G Nartker, TA and SV Rice. (1999) Optical character recognition: An illustrated guide to the frontier. *Proceedings of SPIE*

- Simard, P.Y. Szeliski, R.. Benaloh J,. Couvreur J, and Calinov I..(2003) Using character recognition and segmentation to tell computer from humans. *Document Analysis and Recognition*, 1:418

- Mori and Malik J. (2003) Recognizing objects in adversarial clutter: Breaking a visual CAPTCHA. 2003

- Chellapilla K. and Simard P.. (2004) Using machine learning to break visual human interaction proofs (HIPs). *Advances in Neural Information Processing Systems*, 17

- Wilkins. J (2009) Strong CAPTCHA guidelines v1. 2.

- Grainger J and Jacobs AM. (1996) Orthographic processing in visual word recognition: A multiple read-out model. Psychological review

- Frost. R (1998) Toward a strong phonological theory of visual word recognition: True issues and false trails. Psychological Bulletin

- Lavrenko, V Rath, TM and Manmatha R (2004) . Holistic word recognition for handwritten historical documents. First International Workshop on Document Image Analysis for Libraries, 2004. Proceedings, 2004.

- CAPTCHA

1.5. ΤΕΧΝΙΚΑ ΣΤΟΙΧΕΙΑ

Τα CAPTCHAs θα πρέπει ουσιαστικά να καλύπτουν τρεις βασικές αρχές (Chew and Tygar, 2004). Αυτές είναι:

- Να είναι εύκολο στη λύση για τον άνθρωπο
- Να είναι εύκολο για τη μηχανή να το τεστάρει και να το βαθμολογήσει
- Δύσκολο για ένα ρομπότ και ένα λογισμικό να το σπάσουν. Ο μοναδικός μηχανισμός που θα πρέπει να σπάει το CAPTCHA θα είναι αυτό που το δημιουργεί.

Σύμφωνα με τον Chellapilla et al. (2005) στόχος του CAPTCHA είναι να επιτυγχάνεται ποσοστό αποτυχία μικρότερο από 0.01% και το αντίστοιχο ποσοστό επιτυχίας για τον ανθρώπινο παράγοντα να είναι 90%. Ένα CAPTCHA που βασίζεται σε κείμενο θα πρέπει να εκμεταλλεύεται τα προβλήματα της τεχνιτής νοημοσύνης και της αναγνώρισης χαρακτήρων. Γι αυτό το λόγο θα πρέπει να είναι ανθεκτικό στην κατάτμηση των λέξεων (segmentation resistant) στους επιμέρους χαρακτήρες. Για την ώρα δεν υπάρχει τρόπο να αποδειχθεί η ανθεκτικότητα ενός CAPTCHA στην κατάτμηση των λέξεων παρά μόνο μέσα από την εφαρμογή εμπειρικών μεθόδων μέσα από OCR. Οι λέξεις που θα χρησιμοποιούνται θα πρέπει να είναι πραγματικές λέξεις από το υφιστάμενο λεξιλόγιο.

Ποια όμως είναι τα βασικά χαρακτηριστικά που θα πρέπει να εκπληρώνει ένα CAPTCHA; Συγκεκριμένα θα πρέπει να καλύπτονται στόχοι όπως η χρηστικότητα, η ασφάλεια και η πρακτικότητα:

- Χρηστικότητα: η χρηστικότητα αφορά στο βαθμό δυσκολίας να επιλύονται τα CAPTCHA από τον άνθρωπο. Σημαντικός παράγοντας για την επίτευξη του στόχου είναι και ο χρόνος. Όσο μεγαλύτερος είναι ο βαθμός δυσκολίας για την επίλυση από το χρήστη τόσο δύσχρηστο θεωρείται το CAPTCHA.
- Ασφάλεια: η ασφάλεια αφορά στο βαθμό δυσκολίας που τίθεται στο εκάστοτε λογισμικό να επιλύσει το γρίφο.
- Πρακτικότητα: η πρακτικότητα αφορά στο βαθμό που είναι διατεθειμένος ο χρήστης να χρησιμοποιήσει την εφαρμογή. Παραδείγματος χάριν θα πρέπει να

είναι ικανός να επιλύσει το CAPTCHA τόσο από έναν τυπικό browser όσο και από ένα κινητό τηλέφωνο.

Σύμφωνα με τον Nielsen (2003) η χρηστικότητα διακρίνεται σε πέντε στοιχεία ποιότητας:

- Την εκμάθηση: Πόσο εύκολα ένας χρήστης μπορεί να ολοκληρώσει τις βασικές λειτουργίες την πρώτη φορά που έρχεται σε επαφή με την εφαρμογή.
- Την αποτελεσματικότητα: Από τη στιγμή που ο χρήστης μάθει την εφαρμογή σε πόσο χρονικό διάστημα μπορεί να ολοκληρώσει μία διαδικασία.
- Την απομνημόνευση: Μετά από τη διακοπή της χρήσης μιας εφαρμογής από το χρήστη κατά πόσο μπορεί να ανταποκριθεί στις απαιτήσεις της μετά από την επαναχρησιμοποίησή της.
- Λάθη: Σε ποιο βαθμό κάνει λάθη ο χρήστης και πόσο σημαντικά είναι αυτά. Πόσο γρήγορα μπορεί ο χρήστης κατανοήσει τα λάθη που έχει κάνει;
- Ικανοποίηση: Πόσο ευχάριστος είναι ο σχεδιασμός;

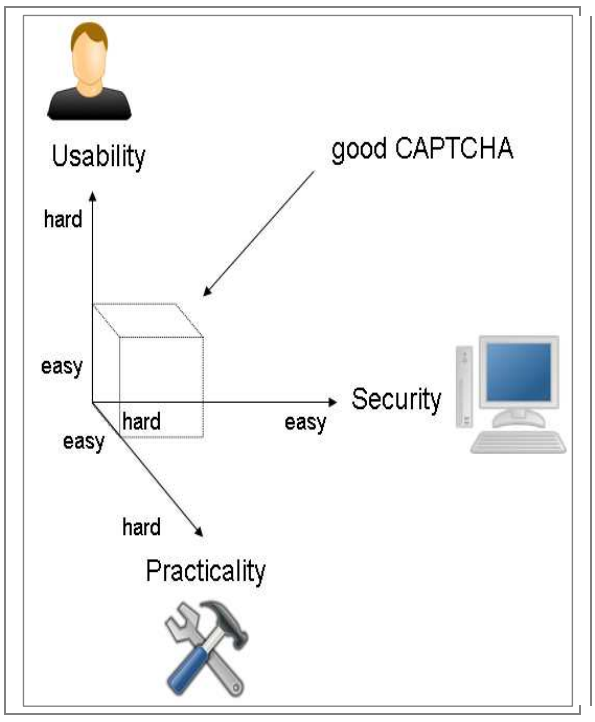
Από όλα τα παραπάνω τη μεγαλύτερη βαρύτητα κερδίζει όπως είναι λογικό ο βαθμός κατανόησης της εφαρμογής. Φυσικά η εφαρμογή έχει να επιδείξει υψηλό βαθμό εκμάθησης και απομνημόνευσης. Από την άλλη θα πρέπει να τεθούν στο μικροσκόπιο ζητήματα όπως:

- Η ακρίβεια: με πόση ακρίβεια μπορούν οι χρήστες να περάσουν τη δοκιμασία του CAPTCHA; Για παράδειγμα πόσες προσπάθειες απαιτείται να κάνουν οι χρήστες για να ολοκληρώσουν επιτυχώς τη δοκιμασία;
- Ο χρόνος απόκρισης: πόση ώρα θα χρειαστεί για να ολοκληρώσουν τη δοκιμασία οι χρήστες;
- Η αντιλαμβανόμενη δυσκολία - ικανοποίηση: πόσο δύσκολη θεωρούν οι χρήστες ότι είναι η εφαρμογή; Οι χρήστες θεωρούν ότι είναι ικανοποιημένοι από τη χρήση της;

Τα παραπάνω κριτήρια μπορούν να προσδιορίσουν ποσοτικά την ποιότητα ευχρηστίας της εφαρμογής CAPTCHA. Μετά από την ανάλογη μελέτη οι ενδιαφερόμενοι σχεδιαστές μπορούν να αναγνώσουν τα αποτελέσματα της έρευνας με ακρίβεια και να βελτιώσουν την εμπειρία του χρήστη.

Πέρα από τα προαναφερθέντα κριτήρια η βιβλιογραφία γίνεται πιο συγκεκριμένη όσον αφορά στα κριτήρια ευχρηστίας που μπορούν να εφαρμοστούν σχετικά με το CAPTCHA. Αυτά είναι:

- Παραμόρφωση. Η διάσταση αυτή εξετάζει τη μορφή στρεβλώσεις που εφαρμόζεται από ένα CAPTCHA και τις επιπτώσεις της στην χρηστικότητα.
- Περιεχόμενο. Η διάσταση αυτή εξετάζει τα περιεχόμενα ενσωματωμένα σε CAPTCHA προκλήσεις (ή διαγνωστικές εξετάσεις) και οι επιπτώσεις τους στη χρηστικότητα. Για παράδειγμα, πώς θα πρέπει να είναι το περιεχόμενο οργανωμένο, και είναι το περιεχόμενο κατάλληλο;
- Η παρουσίαση. Η διάσταση αυτή εξετάζει τον τρόπο με τον οποίο παρουσιάζονται οι CAPTCHA προκλήσεις που και οι επιπτώσεις της στη χρηστικότητα.



Εικόνα 2-4: Χαρακτηριστικά ενός καλού CAPTCHA

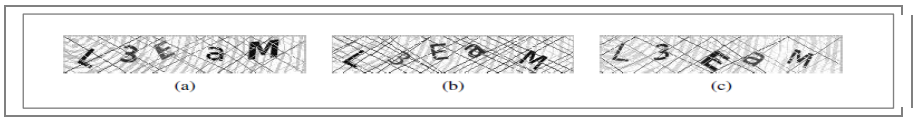
Η ιδανική περίπτωση ανάπτυξη ενός CAPTCHA θα ήταν να αποτελέσει σημείο υψηλού βαθμού ευχρηστίας από τον χρήστη και ταυτόχρονα υψηλή ασφάλεια από τις επιθέσεις. Από την άλλη πλευρά η βάση δεδομένων των εφαρμογών δεν θα πρέπει να

είναι δημόσια καθώς θα αποτελέσει σημείο εκμάθησης για κάθε προσπάθεια μηχανικής εκμάθησης.

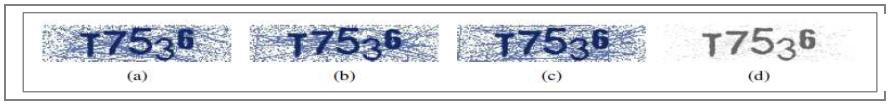
Πότε όμως μπορεί να θεωρηθεί ότι μία εφαρμογή CAPTCHA έχει σπάσει; Ο βαθμός ασφάλεια μίας εφαρμογής προσδιορίζεται μέσα από το ποσοστό των περιπτώσεων που σπάνε. Συνήθως αν οι προσπάθειες επίλυσης των γρίφων είναι επιτυχημένες πάνω από το 5% τότε θεωρείται η εφαρμογή είναι αδύναμη. Πολλές επιθέσεις σε CAPTCHA εντάσσουν και τον ανθρώπινο παράγοντα σε συνδυασμό με τη μηχανική προσπάθεια. Πολλά άτομα επιλύουν CAPTCHA με τη λογική να δημιουργήσουν μία βάση εκμάθησης των μηχανικών τεχνικών.

1.6. ΒΑΣΙΚΆ ΣΤΟΙΧΕΊΑ ΣΧΕΔΙΑΣΜΟΥ

Έχοντας δώσει τα βασικά ιστορικά στοιχεία και τη χρησιμότητα του CAPTCHA δόκιμο είναι να δοθούν και τα βασικά στοιχεία σχεδιασμού ενός τέτοιου μηχανισμού. Όταν ένας ενδιαφερόμενος επιθυμεί να αναπτύξει μία εφαρμογή CAPTCHA κειμένου θα βρει εξαιρετικά ελκυστικό να βασιστεί σε ένα λεξικό για να τροφοδοτήσει με υλικό το μηχανισμό του. Σίγουρα η τακτική αυτή μοιάζει ιδιαίτερα ελκυστική αφού το υλικό που θα τροφοδοτήσει τον μηχανισμό είναι πολύ εύκολο να συλλεχθεί, αλλά σίγουρα η τακτική αυτή θα αποφέρει πολλά μειονεκτήματα. Το πιο προφανές είναι ότι διευκολύνεται κάθε είδους επίθεση καθώς οι μηχανισμοί συνήθως εκπαιδεύονται με βάση τα λεξικά. Αυτό περιορίζει το χώρο που έχει το άτομο που θα αναπτύξει την εφαρμογή να προκαλέσει τον επιτιθέμενο, καθώς όλες οι πιθανές απαντήσεις θα είναι προϋπολογισμένες. Ταυτόχρονα οι εφαρμογές κειμένου προϋποθέτουν την γνώση της γλώσσας από το άτομο που προσπαθεί να το λύσει. Αυτό το μειονέκτημα ευχρηστίας μπορεί να παρακαμφθεί μέσα από την ανοχή ορισμένων λαθών από την πλευρά του χρήστη. Το άτομο που θα αναπτύξει μία εφαρμογή κειμένου θα πρέπει να αφήνει κάποιο χώρο ανάμεσα στους χαρακτήρες προκειμένου να διευκολύνει τη χρήση από τον άνθρωπο. Στις εικόνες που ακολουθούν μπορούμε να παρακολουθήσουμε ένα CAPTCHA με διαφορετικές παρουσιάσεις προς τον τελικό χρήστη.



Εικόνα 2-5 Εικόνες Captcha 1/2



Εικόνα 2-6 Εικόνες Captcha 2/2

Η άσκοπη χρήση διαφορετικών χρωμάτων δεν μπορεί να ωφελήσει ούτε τον τελικό χρήστη ούτε και τους επιθετικούς μηχανισμούς. Το βασικό πρόβλημα με αυτή την τεχνική είναι ότι κάθε αλλαγή χρώματος μπορεί να αποτελέσει σημείο αναφοράς για τον επιτιθέμενο μηχανισμό για να ξεκινήσει την τμηματοποίηση της λέξης. Για παράδειγμα η δημιουργία μίας ακολουθίας διαφορετικών χαρακτήρων με διαφορετικό χρώμα απλουστεύει πολύ το βαθμό δυσκολίας αφού το εκάστοτε χρώμα αποτελεί σημείο αναγνώρισης. Από την άλλη επικρατεί η εντύπωση ότι η υπερφόρτωση με χρώμα του background της εικόνας μπορεί να λειτουργήσει αποτρεπτικά. Παρόλα αυτά υπάρχουν αναρίθμητα παραδείγματα επιθέσεων σε CAPTCHA που είχαν τέτοιου είδους παραμορφώσεις. Ο θόρυβος αυτός σχετίζεται με τη χρήση του χρώματος στο σώμα της λέξης. Η υπερφόρτωση του πλαισίου στο οποίο έχει τοποθετηθεί η λέξη αυξάνει την πιθανότητα να θεωρηθεί δύσχρηστη η εφαρμογή. Οι Chellapilla και συν. (2005) θεωρούν ότι οι σχεδιαστές των συστημάτων θα πρέπει να δώσουν βάση τόσο στην αντιμετώπιση της τμηματοποίησης των λέξεων όσο και στην αναγνώριση τους.

Οι σχεδιαστές θα πρέπει να εντάξουν στη διαδικασία σχεδιασμού του CAPTCHA και τον ανθρώπινο παράγοντα με τις ιδιομορφίες του. Πριν από τον τελικό σχεδιασμό καλό θα ήταν να εξεταστεί η ευχρηστία και η αξιοπιστία του CAPTCHA από τον άνθρωπο. Από τη στιγμή που ο άνθρωπος μπορεί να επιλύσει θέματα μέχρι ένα ποσοστό, το ποσοστό αυτό μπορεί να γενικευτεί και να εξαχθούν ασφαλή συμπεράσματα.

Ο επιτιθέμενος μηχανισμός στοχεύει στην αδυναμία του CAPTCHA. Ένα πολύ σημαντικό μειονέκτημα σε εφαρμογές του CAPTCHA είναι η επανάληψη των προβλημάτων που τίθενται στο χρήστη. Η επανάληψη μιας λέξης αποτελεί πύλη εισόδου για τον εκάστοτε κακόβουλο μηχανισμό. Το σύστημα θα πρέπει να καταγράφει τα προβλήματα που επιλύονται και να μην επαναλαμβάνει την εμφάνισή τους.

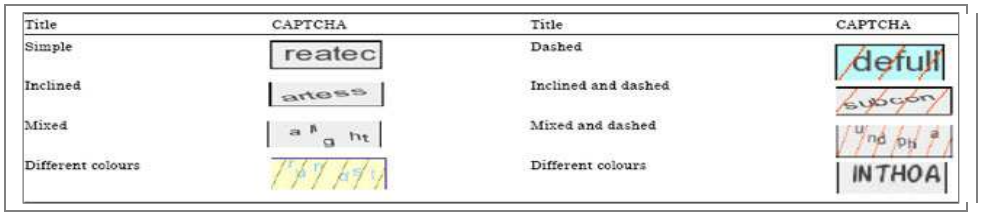
Εικόνα 2-7 Η εξέλιξη του CAPTCHA

Μία άλλη βασική πτυχή του ζητήματος που πρέπει να μελετηθεί είναι η μορφή που θα λαμβάνει το CAPTCHA κάθε φορά που θα εμφανίζεται στο χρήστη. Ο βαθμός και το είδος της παραμόρφωσης και το είδος της γραμματοσειράς και των συνδυασμών θα πρέπει να διαφέρουν. Όπως διαφαίνεται και από την παραπάνω εικόνα οι συνδυασμοί και οι παραμορφώσεις που μπορούν να γίνουν είναι πολλοί. Κάτι που πρέπει να γίνει κατανοητό είναι ότι δεν πρέπει οι επιτιθέμενοι να προσαρμοστούν στη λογική του μηχανισμού, αλλά αντίθετα ο μηχανισμός να ανταποκριθεί στη λογική των επιτιθέμενων. Το σύστημα θα πρέπει να παρατηρεί τους περίεργους χρήστες και να αφομοιώνει τη συμπεριφορά τους προκειμένου να ανταποκριθούν στις απαιτήσεις τους. Η καταγραφή των IP είναι μία πολύ σημαντική μέθοδος αποτροπής επαναλαμβανόμενων επιθέσεων.

1.7. ΕΙΔΗ CAPTCHA

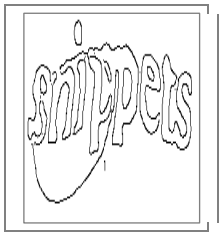
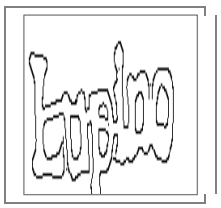
1.7.1. Κειμενικό CAPTCHA

Σύμφωνα με τους Yan και Ahmad 2008 τα κειμενικά CAPTCHA τυπικά βασίζονται στην εξεζητημένη παραμόρφωση εικόνων, λέξεων καθιστώντας τις μη αναγνωρίσιμες στην πλειοψηφία του OCR λογισμικού, αλλά την ίδια ώρα αναγνωρίσιμο από το μέσο χρήστη του διαδικτύου. Η συγκεκριμένη μορφή CAPTCHA ήταν και η πρώτη που έκανε την εμφάνισή της. Το κειμενικό CAPTCHA σχεδιάστηκε για να αντλεί το υλικό που θα παρατίθεται στο χρήστη μέσα από ένα πλούσιο σύνολο λέξεων. Στις λέξεις αυτές ένας μηχανισμός θα λειτουργούσε παρεμβατικά προκειμένου να παραμορφώσει την εμφάνισή τους και με αυτό τον τρόπο να αυξήσει τη δυσκολία επίλυσης του γρίφου από ένα λογισμικό. Η συγκεκριμένη μορφή CAPTCHA είναι η πιο διαδεδομένη και μερικές από τις εκφάνσεις που μπορεί να πάρει παρουσιάζονται στην παρακάτω εικόνα (Zhu και συν., 2010).



Εικόνα 2-8 Είδη κειμενικού CAPTCHA

Ο Milde (2010) αναφέρει ότι έχει αναλύσει ένα σύνολο 130 λέξεων κατά την περίοδο Απριλίου του 2010. Σύμφωνα με τον Mide (2010) οι μορφές παραμόρφωσης που μπορούν να εφαρμοστούν πάνω στις λέξεις είναι οι παρακάτω.

Κατηγορία	%	Παραδείγματα	Περιγραφή
Μικρή παραμόρφωση	44,6%		Ένα μεγάλο μέρος της έλλειψης είναι ξεκάθαρα φανερή και υπάρχει αρκετός χώρος για τους χαρακτήρες.
Μέτρια παραμόρφωση	43,9%		Η έλλειψη διέρχεται ανάμεσα σε ένα μεγάλο μέρος της λέξης ή τουλάχιστον καλύπτει ένα μέρος της και η παραμόρφωση είναι μεγαλύτερη.
Ισχυρή παραμόρφωση	6,9%		Η έλλειψη δεν εφαρμόζεται απόλυτα στη ροή της λέξης και αυτό έχει ως συνέπεια την ευρύτερη παραμόρφωσή της.
Άλλη	4,6%		Το CAPTCHA δεν έχουν επιπλέον στρέβλωση στην μορφή μιας έλλειψης (είναι ένα διαφορετικό είδος CAPTCHA).

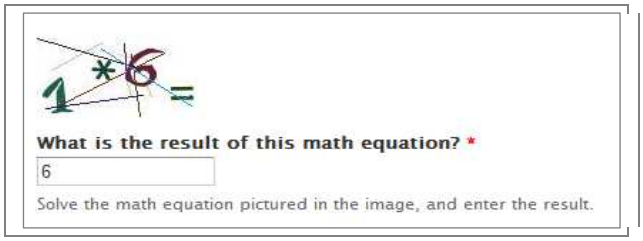
1.7.2. Σχήματα βασισμένα στον ήχο

Τα συγκεκριμένα συστήματα απαιτούν την τυπική επίλυση αναγνώρισης ενός ήχου (Yan και Ahmad 2008). Η συγκεκριμένη μορφή CAPTCHA που βασίζεται στον ήχο είχε ως πρωταρχικό σκοπό να απευθυνθεί σε άτομα που είχαν προβλήματα όρασης. Το άτομο που θέλει να έχει πρόσβαση σε μία συγκεκραμένη πηγή θα πρέπει σε πρώτο στάδιο να αναγνωρίσει το γρίφο που του παρουσιάζεται. Επομένως τα άτομα που έχουν προβλήματα όρασης δεν μπορούσαν να έχουν πρόσβαση. Ο χρήστης ακούει έναν ήχο τον οποίο καλείται να γράψει στο κενό πεδίο. Οι Bigham και Cavender (2009) θεωρούν ότι η μέθοδος του ηχητικού CAPTCHA είναι χρονοβόρα και δεν έχουν υψηλό βαθμό ασφαλείας (Gupta και συν., 2009).

Εικόνα 2-9 Το audio CAPTCHA

1.7.3. CAPTCHA μαθηματικής πράξης

Τα CAPTCHA μαθηματικής πράξης είναι το επόμενο βήμα σε αυτό που αναφέραμε ως κειμενικό CAPTCHA. Σύμφωνα με τη συγκεκριμένη μορφή ο χρήστης όχι μόνο θα πρέπει να αναγνωρίσει τους χαρακτήρες που του παρουσιάζονται, αλλά θα πρέπει και να ολοκληρώσει επιτυχώς και μία μαθηματική πράξη που εμφανίζεται.

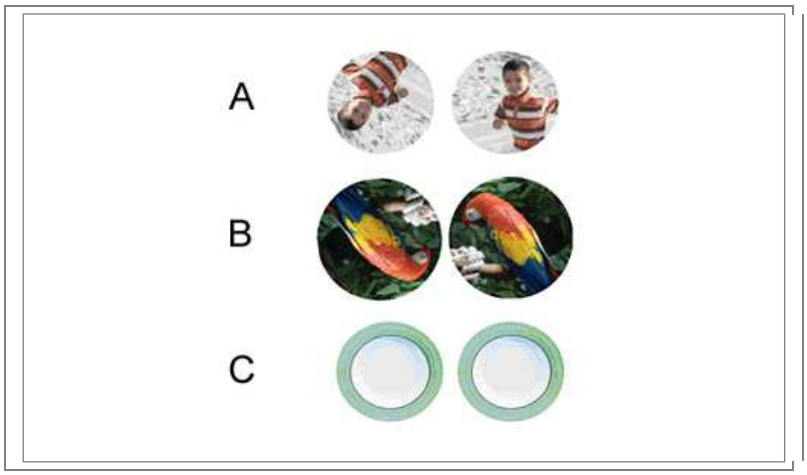


Εικόνα 2-10 Μαθηματικά CAPTCHA

1.7.4. Αναστροφή εικόνας

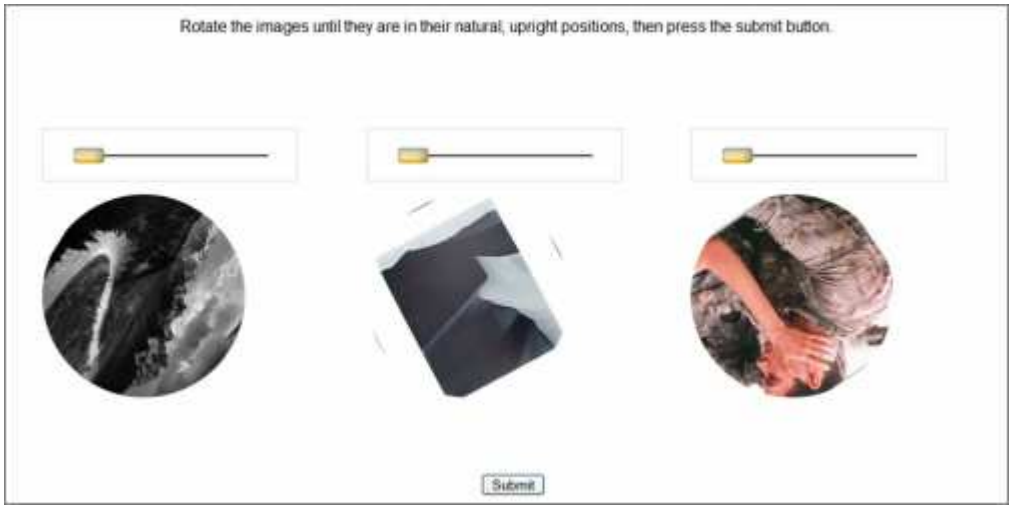
Το σύστημα CAPTCHA που αφορά στην αναστροφή εικόνας βασίζεται στην ικανότητα του ανθρώπου να μπορεί να εντοπίζει τη σωστή φορά της εικόνα και να τη διορθώνει. Σίγουρα για ένα μεγάλο εύρος εικόνων η αυτοματοποιημένη διαδικασία

αναστροφής μπορεί να κατασταθεί εύκολη, αλλά υπάρχουν πάντα εικόνες που δεν μπορούν να αναγνωριστούν μηχανικά, αφού δεν αποτελούν συνηθισμένες λήψεις. Στην παρακάτω εικόνα μπορούμε να διακρίνουμε τρία ζεύγη εικόνων. Ένα στοιχείο από τα δύο που αποτελούν το ζεύγος είναι λανθασμένο και καλείται ο χρήστης να επιλέξει αυτό που έχει τη λανθασμένη φορά.



Εικόνα 2-11 Είδος CAPTCHA που απαιτεί αναστροφή των εικόνων

Στην περίπτωση που ακολουθεί ο χρήστης καλείται να χρησιμοποιήσει τις μπάρες που βρίσκονται πάνω από κάθε εικόνα προκειμένου να τις φέρει στην σωστή τους φορά.

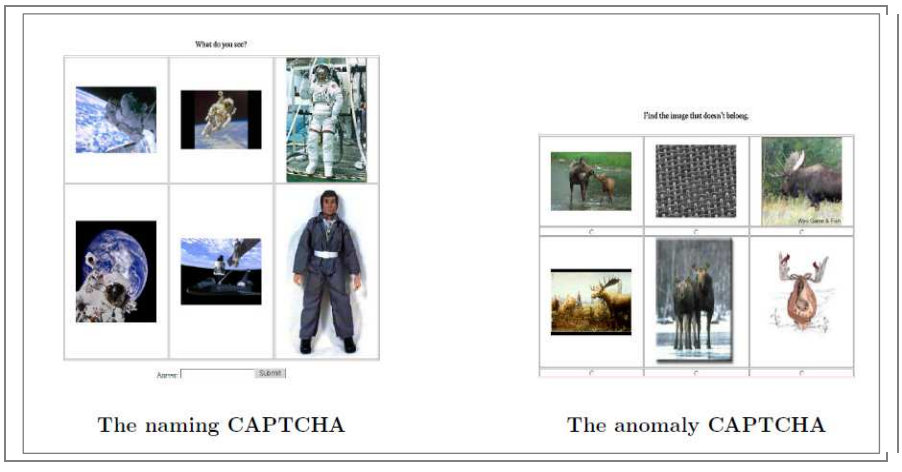


Εικόνα 2-12 Διαμόρφωση σωστής φοράς

1.7.5. Αναγνώριση εικόνας

Η μέθοδος της αναγνώρισης εικόνας έχει να παρουσιάσει διάφορες εκδοχές. Συγκριμένα:

- **Η ονοματοθεσία των εικόνων**: Σύμφωνα με αυτή τη μέθοδο ο χρήστης παρακολουθεί έξι εικόνες. Κάτω από τις εικόνες υπάρχει ένα κενό πλαίσιο στο οποίο καλείται να συμπληρώσει την έννοια που απεικονίζεται στην εικόνα. Στο παράδειγμα της εικόνα που ακολουθεί ο χρήστης θα πρέπει να συμπληρώσει τη λέξη «Αστροναύτης».
- **Η διάκριση των εικόνων**: Ο συγκεκριμένος μηχανισμός CAPTCHA παρουσιάζει δύο σύνολα εικόνων για τον χρήστη. Κάθε το σετ περιέχει τρία εικόνες του ίδιου θέματος με την ίδια πιθανότητα να έχουν είτε το ίδιο θέμα ή όχι. Ο χρήστης πρέπει να καθορίσει κατά πόσον έχουν το ίδιο αντικείμενο ή όχι προκειμένου να περάσουν στον επόμενο γύρω.
- **Ο προσδιορισμός ανωμαλιών**: Σύμφωνα με το συγκεκριμένο μοντέλο παρουσιάζεται στον χρήστη ένα σύνολο εικόνων οι οποίες συνδέονται μεταξύ τους αφού παρουσιάζουν το ίδιο θέμα. Στόχος του χρήστη είναι να προσδιορίσει την εικόνα που δεν έχει το ίδιο θέμα με τις υπόλοιπες.



Εικόνα 2-13: Η αναγνώριση των εικόνων και των ανωμαλιών

Μία παραλλαγή του συστήματος ονοματοθεσία των εικόνων αποτελεί το ESP – Pix CAPTCHA. Σύμφωνα με το ESP – Pix CAPTCHA ο χρήστης θα πρέπει να

αναγνωρίσει το κοινό στοιχείο των εικόνων και στη συνέχεια θα πρέπει να επιλέξει ανάμεσα σε διαθέσιμες απαντήσεις.

Εικόνα 2-14: Το ESP – Pix CAPTCHA

Στον πίνακα που ακολουθεί μπορούμε να παρατηρήσουμε στο σύνολό τους τα είδη των CAPTCHA που υπάρχουν:

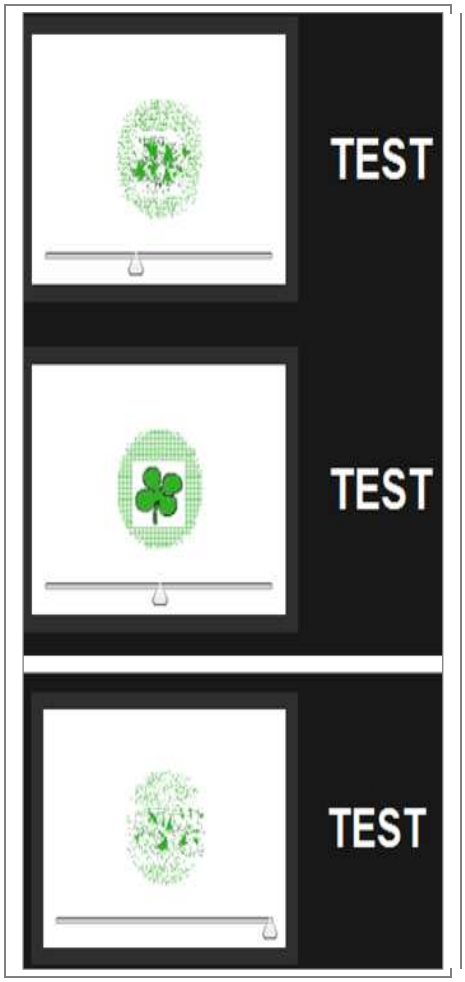
Τύπος CAPTCHA	Τίτλος
Κειμενικά CAPTCHA	Πρώιμα CAPTCHA
	Βελτιωμένα CAPTCHA
	Μοντέρνα CAPTCHA
	Κεινούμενα CAPTCHA
	ASCII CAPTCHA
	Αντίστροφα CAPTCHA
	CAPTCHA επίλυσης προβλημάτων
CAPTCHA που βασίζονται στην εικόνα	CAPTCHA αναγνώρισης εικόνας
	3D CAPTCHA

Ηχητικά CAPTCHA	Ήχος
CAPTCHA video	NuCAPTCHA

1.7.6. 3d CAPTCHA

Ο Rolko (2010) σε μία ενδιαφέρουσα μελέτη του ανέλυσε τη λειτουργία ενός 3d CAPTCHA και τον τρόπο δημιουργίας του. Η βασική λογική του συγκεκριμένου τύπου CAPTCHA είναι να ανακαλύψει ο χρήστης την ορθή θέση του σχήματος μέσα από μια διαδικασία περιστροφής του. Η τρισδιάστατη εικόνα δημιουργείται από μία απλή δισδιάστατη εικόνα, η οποία διαιρείται σε μικρά κομμάτια τα οποία κατανέμονται στο χώρο. Ο χρήστης μέσα από την κύλιση μιας μπάρας προσπαθεί να βρει τη σωστή προοπτική που θα παρουσιάσει ξεκάθαρα την εικόνα μπροστά του. Η τρισδιάστατη μορφή της εικόνας μπορεί να γίνει ορατή μόνο από ένα συγκεκριμένο σημείο. Η παρουσίαση της εικόνας είναι κατανοητή από τον τελικό χρήστη προκειμένου να μπορέσει να λύσει το γρίφο. Θεωρείται δεδομένο ότι οι συμβατικές τεχνολογίες δεν μπορούν να διασπάσουν το συγκεκριμένο επίπεδο ασφάλειας. Τα βήματα δημιουργίας ενός 3d CAPTCHA συνοψίζονται στα εξής:

- Επιλογή της επιθυμητής εικόνας που θα αποτελέσει την εικόνα του γρίφου.
- Εφαρμογή πρόσθετων φίλτρων πάνω στη δισδιάστατη εικόνα
- Διαίρεση της εφαρμογής σε βασικά συστατικά στοιχεία
- Διασκορπισμός των κομματιών πάνω στο χώρο.
- Δημιουργία ενός 3D μοντέλου
- Ενσωμάτωση της εφαρμογής στον client



Εικόνα 2-15 Το 3d CAPTCHA

1.8. Δημοφιλή CAPTCHA scripts

Στην παρούσα ενότητα θα παρουσιαστούν τα πιο δημοφιλή scripts για ενσωμάτωση CAPTCHA που είναι διαθέσιμα και δωρεάν. Συγκεκριμένα είναι:

1.8.1. reCAPTCHA

Το reCAPTCHA (reCAPTCHA.net) είναι ίσως το επιτυχημένο από τα συστήματα CAPTCHA. Όπως αναφέρθηκε η δυναμική της οφείλεται στον αναρίθμητο πλούτο λέξεων στον οποίο βασίζεται και μπορεί να παρουσιάσει. Κάθε φορά που ένα σύστημα οπτικής αναγνώρισης χαρακτήρων διενεργεί επίθεση στο σύστημα αυτό μαθαίνει από τη τακτική του και καταχωρεί στο ιστορικό την επίθεση αυτή. Το θετικό στοιχείο του reCAPTCHA είναι ότι αποτελεί μία κεντρικοποιημένη υπηρεσία και όχι ένα πρόγραμμα που μπορεί κανείς να το προμηθευτεί ή να το κατεβάσει από το διαδίκτυο. Ένα άλλο σημαντικό στοιχείο είναι ότι δεν υπάρχει μια σταθερή και

δεδομένη εκδοχή του μηχανισμού, αφού αποτελεί ένα πολύ δυναμικό σύστημα που αλλάζει δραστικά στο πέρασμα του χρόνου. Οι σπασμωδικές ενδείξεις σπασίματος ενός συστήματος reCAPTCHA είναι απλά αποδείξεις του μεγάλου βαθμού ασφαλείας που παρέχει το σύστημα. Ταυτόχρονα είναι πολύ δύσκολο να αποδειχθούν οι επιτυχίες απέναντι στο σύστημα καθώς θα πρέπει να έχουν μεγάλο βαθμό επαναληψιμότητας.

Εικόνα 2-16 reCAPTCHA

Ο Wilkins (2009) σημειώνει ορισμένες μεθόδους οι οποίες μπορούν να φανούν χρήσιμες σε προσπάθειες επίλυσης ενός συστήματος reCAPTCHA αναφέροντας ότι χρειάζονται μορφολογικοί μηχανισμοί που μπορούν να διαγράψουν ορισμένα pixels από τις ακμές των γραμμάτων ενώ ταυτόχρονα θα πρέπει να προσθέσουν κάποια άλλα. Η τελευταία εκδοχή του reCAPTCHA βγήκε στην κυκλοφορία τον Απρίλιο του 2010 και η οποία τυχαία αλλάζει τους χρωματισμούς των χαρακτήρων. Μέσα από πειραματικές διαδικασίες έχει παρατηρηθεί ότι μόλις το 5% των παρουσιασθέντων προβλημάτων μπορεί να λυθεί.

Ένα ζήτημα σχετικά με την εξέλιξη reCAPTCHA είναι ότι αν οι τελευταίες βελτιστοποιήσεις βελτιώνουν σημαντικά την ασφάλεια. Αν και δεν είμαστε ενήμεροι για τυχόν επιθέσεις σε αυτό το στάδιο, υπάρχουν μερικές ιδιότητες που κάνουν αυτό το είδος της στρέβλωσης πιθανώς ευάλωτο:

- Υπάρχει μικρή διακύμανση στο γενικό σχήμα του αντικειμένου. Είναι πάντα καλά-στρογγυλευμένο και μπορεί να προσεγγιστεί χονδρικά με μία έλλειψη.
- Οι ακμές του αντικειμένου παραμόρφωσης είναι πολύ "καθαρές", είναι συνεπώς εύκολο να ανιχνεύσει αυτά ένα πρότυπο αλγορίθμων. Οι ακμές έχουν χαμηλή καμπυλότητα σε σύγκριση με τις καμπύλες των χαρακτήρων και συχνά εκτείνονται σε περιοχές όπου αυτά ξεχωρίζουν.

1.8.2. Asira

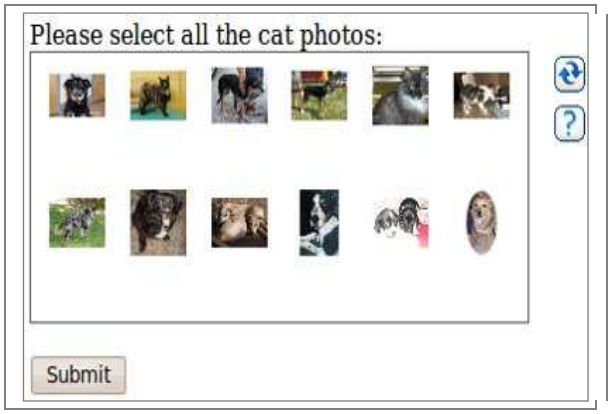
Οι Elson και συν. (2007) αναφέρουν την ύπαρξη ενός νέου δυναμικού και ελπιδοφόρου εργαλείου με ενθαρρυντικά αποτελέσματα. Το Asira (Animal Species Image Recognition for Restricting Access) της Microsoft Research είναι ένα ανθρώπινος διαδραστικός μηχανισμός που παρέχει μια βιώσιμη εναλλακτική λύση και όχι την κλασική CAPTCHAs. Το Asira δεν αποτελεί προϊόν παρθενογέννεσης, αντίθετα αποτελεί μετεξέλιξη των Frozen Bear HotCAPTCHA, του Carnegie Mellon PIX CAPTCHA και του Oli Warner KittenAuth. Παρόλα αυτά η βασική διαφορά με τα υπόλοιπα εργαλεία είναι το γεγονός ότι χρησιμοποιεί μία βάση 3 εκατ. εικόνων από το Petfinder.com, ενώ τα υπόλοιπα εργαλεία χρησιμοποιούσαν αρκετά περιορισμένες βάσεις. Το βασικό μειονέκτημα των προηγούμενων εργαλείων είναι ότι ο εκάστοτε κακόβουλος μπορεί να έχει πρόσβαση στο υλικό που χρησιμοποιούν ως βάση. Μερικά είδη CAPTCHA μπορεί να είναι ιδιαίτερα προσβλητικά για το χρήστη, αφού μπορεί να χρησιμοποιούν ακόμα και προσβλητικό υλικό για το διαχωρισμό μηχανής και ανθρώπου. Για παράδειγμα το Hot CAPTCHA χρησιμοποιεί φωτογραφίες ανθρώπων και καλεί τον χρήστη να αξιολογήσει το βαθμό σεξουαλικότητας που αποπνέει μία φωτογραφία. Στη φωτογραφία που ακολουθεί παρουσιάζεται το λογότυπο του Asira.

Εικόνα 2-17 Λογότυπο του Asirra

Από την άλλη πλευρά το Asira αυξάνει το βαθμό δυσκολία στον εκάστοτε επιτιθέμενο αφού βασίζεται σε μία πολύ ισχυρή βάση. Επιπλέον όλη η βάση του

Petfinder.com δε είναι διαθέσιμη στο κοινό. Αλλά ταυτόχρονα είναι στη Microsoft Research.

Όσον αφορά στη λειτουργία του μηχανισμού είναι εξαιρετικά απλός και μοιάζει λίγο έως πολύ με του παραπάνω. Ο χρήστης καλείται να επιλέξει από ένα σύνολο εικόνων που απεικονίζουν ζώα αυτό που του υποδεικνύει η ερώτηση. Για παράδειγμα στη δοκιμασία που παρουσιάζεται στην παρακάτω εικόνα, ο χρήστης καλείται να επιλέξει τη φωτογραφία που απεικονίζει τη γάτα.



Εικόνα 2-18 Quiz τύπου Asirra

Είναι εύκολο να προσθέσει κανείς ένα HIP Asirra στο web site του. Η Microsoft Research παρέχει δωρεάν την υπηρεσία web αν και προειδοποιεί ότι η κατάστασή της είναι υπό δοκιμή. Το API της μπορεί να είναι ασταθές. Το Asirra αποτελείται από:

- Ένα JavaScript client που προστίθεται στην ιστοσελίδα μέσα σε μια φόρμα. Ο κωδικός που δίνεται από το Asirra θα προσθέσει μια πρόκληση στην ιστοσελίδα. Εάν η πρόκληση λυθεί σωστά, ο κώδικας πελάτης παίρνει ένα εισιτήριο Asirra από τον server.
- Μια υπηρεσία web της Microsoft Research η οποία θα ελέγχει την εγκυρότητα του εισιτηρίου.

Το JavaScript λειτουργεί σε όλους τους κύριους φυλλομετρητές και έχει ελεγχθεί σε IE6, IE7, Firefox 2, Safari 2.0.4, Opera 8.54 και Opera 9.23

1.8.3. Securimage

Το Securimage είναι ένα script ελεύθερο για τη δημιουργία περίπλοκων εικόνων και κωδικών CAPTCHA προκειμένου να προστατεύσει το χρήστη από spam μηνύματα. Ο κάθε σχεδιαστής μπορεί εύκολα να εντάξει το script στην εφαρμογή του αρκεί βέβαια να έχει εγκαταστήσει PHP.

Εικόνα 2-19: Λογότυπο Securimage

Τα βασικά χαρακτηριστικά του Securimage είναι:

- Μπορεί να παρουσιάσει μία εικόνα με μόλις 3 γραμμές κώδικα
- Μπορεί να επικυρώσει τις καταχωρημένες προσπάθειες με μόλις 6 γραμμές κώδικα
- Υποστήριξη γραμματοσειράς TTF
- Ρυθμιζόμενο εύρος κώδικα
- Εύκολη προσθήκη εικόνων
- Δίνεται η δυνατότητα παραμόρφωσης των εικόνων
- Μπορούν να προστεθούν και ηχητικά αρχεία
- Μπορούν να χρησιμοποιηθούν αλφαριθμητικά CAPTCHAs

1.8.4. WebSpamProtect

Το WebSpamProtect επιτρέπει την προσθήκη εικόνα επαλήθευσης (CAPTCHA) στην ιστοσελίδα και προστατεύσει τις φόρμες από το spam ρομπότ. Για την προστασία της φόρμα, το σύστημα απαιτεί την εγκατάσταση ενός μικρού κομματιού κώδικα στην ιστοσελίδα. Ο κώδικας απαιτεί web PHP υποστήριξη, Perl, ASP ή ASP.NET. Το βασικό πρόγραμμα είναι δωρεάν, αλλά μπορείτε να αγοράσει ο κάθε ενδιαφερόμενος ένα Premium, Advanced και Professional με μια μικρή δαπάνη.

Εικόνα 2-20: WebSpamProtect

Τα βασικά χαρακτηριστικά του WebSpamProtect είναι:

- Είναι συμβατικό με PHP, Perl, ASP και ASP.NET
- Δίνει τη δυνατότητα στον ενδιαφερόμενο να τροποποιήσει τις εικόνες που θα χρησιμοποιήσει
- Δίνει τη δυνατότητα στον ενδιαφερόμενο να τροποποιήσει τις γραμματοσειρές
- Υπάρχει η δυνατότητα δημιουργίας ακουστικού CAPTCHA
- Υπάρχει η δυνατότητα να παρασχεθεί reload κομβίου για το χρήστη σε περίπτωση που θελήσει να επανεκκινήσει την προσπάθειά του εξαιτίας κάποιας δυσκολίας.
- Διαφορετικοί αλγόριθμοι παραγωγής CAPTCHA
- Υπάρχει WordPress plugin
- Δωρεάν υποστήριξη

1.8.5. Text CAPTCHA

Το Text CAPTCHA αποτελεί μία εναλλακτική εφαρμογή CAPTCHA που στηρίζεται κατά κύριο λόγο σε λογικές ερωτήσεις που πρέπει να απαντηθούν από το χρήστη. Συγκεκριμένα οι μορφές των CAPTCHA μπορεί να είναι μία από τις παρακάτω:

- Ποιο από το στομάχι, τη σούπα, το βούτυρο, τη μύτη, το λευκό και το γάλα είναι χρώμα;

- Ποιο από την παρακάτω λίστα είναι το όνομα ανθρώπου: τράπεζα, Λίζα, φρούτο, εσώρουχα;
- Ποια λέξη από τις παρακάτω είναι γραμμένη με κεφαλαία;
- Ποιος από τους παρακάτω αριθμούς είναι ο μικρότερος: 59, 63, 2, 14, 85;
- Ποια από την παρακάτω λίστα αποτελούν μέρη του σώματος: δόντια, μύτη, πόδι, πάπια;
- Ποια από τις παρακάτω λέξεις περιλαμβάνει το γράμμα Y: interjects, cockney, draper, oversees

1.8.6. FreeCap

Το FreeCap αποτελεί ένα ελεύθερο λογισμικό που έχει προκαλέσει αίσθηση στην κοινότητα της πληροφορικής με την εξάπλωση που γνωρίζει καθώς τα χαρακτηριστικά του δικαιολογούν την επιτυχημένη του πορεία. Συγκεκριμένα το FreeCap μπορεί να λειτουργήσει σε όλες τις εκδόσεις των Windows και υποστηρίζει ταυτόχρονα μία σειρά φυλλομετρητών όπως Internet Explorer, Netscape, Mozilla, Trillian, Opera κτλ. Ταυτόχρονα η διεπαφή προσφέρει είναι απόλυτα παραμετροποιήσιμη δίνοντας τεράστια ευελιξία στο σχεδιαστή.

1.8.7. NuCAPTCHA

Το NuCAPTCHA θα λέγαμε ότι είναι το επόμενο βήμα, όσον αφορά στα κειμενικά CAPTCHA. Μέσα από έναν εξαιρετικά ανεπτυγμένο μηχανισμό μπορεί να παρουσιάσει ένα σύνολο διαφορετικών γρίφων προς το χρήστη, ενώ διαθέτει έναν επιπλέον μηχανισμό ο οποίος μπορεί να εντοπίσει ύποπτη δραστηριότητα. Σύμφωνα με το σχεδιασμό του οι χαρακτήρες παρουσιάζονται στο χρήστη κινούνται σε ένα σταθερό μοτίβο. Η ταχύτητα αυτή δεν εμποδίζει την αναγνώριση από τον χρήστη, αλλά είναι ικανή να προκαλέσει προβλήματα σε οποιοδήποτε λογισμικό αναγνώρισης χαρακτήρων. Ταυτόχρονα υπάρχει δυνατότητα επανεκκίνησης του γρίφου στην περίπτωση που δεν γίνεται κατανοητή η μορφή που παρουσιάζεται, ενώ προσφέρεται και βοήθεια στο χρήστη.

Εικόνα 2-21 Το NuCAPTCHA

- **WP CAPTCHA-Free**
- **ProtectWebForm**
- Cryptographp
- Κ.α.

1.9. ΜΕΤΡΗΣΕΙΣ CAPTCHA

Στην παρούσα ενότητα θα παρουσιαστούν οι δύο μετρικές που εφαρμόζουν οι Chew και Tygar (2004) για την αξιολόγηση των εφαρμογών CAPTCHAs. Οι μετρικές στοχεύουν να αξιολογήσουν την αποτελεσματικότητα των εφαρμογών με βάση των αριθμό των επαναλήψεων που απαιτούνται από το χρήστη να ολοκληρώσει τη διαδικασία καθώς και το χρόνο που απαιτείται. Κάθε σχεδιαστής μπορεί να βελτιώσει:

- Τον τύπο της πρόκλησης: ο σχεδιαστή θα πρέπει να επιλέξει ανάμεσα στα υπάρχοντα είδα CAPTCHAs (κείμενο, εικόνα κτλ.)
- Ο αριθμός των εικόνων που παρουσιάζονται: ο σχεδιαστής θα πρέπει να επιλέξει τον αριθμό των εικόνων που θα παρουσιαστούν στην πρόκληση.
- Ο αριθμός των γύρων: ο σχεδιαστής θα πρέπει να προσδιορίσει των αριθμό των γύρων των προκλήσεων στις οποίες θα πρέπει να απαντήσει ο χρήστης και ποιος θα είναι ο βαθμός επιτυχίας.
- Τα μοντέλα επιθέσεων: ο σχεδιαστή θα πρέπει να λάβει υπόψη του τα μέσα που θα χρησιμοποιηθούν από κακόβουλο χρήστη για τη διάσπαση των προκλήσεων.

Κάθε άνθρωπος έχει ειδικά γνωρίσματα και διαφορετικό βαθμό ανοχής ώστε να μπορεί να υπομένει πολλούς γύρους προκλήσεων. Ταυτόχρονα τα κακόβουλο λογισμικό διακρίνεται για την ταχύτητα του σε σχέση με τον άνθρωπο. Σύμφωνα με τον παρακάτω τύπο προσδιορίζεται ότι η ταχύτητα του λογισμικού είναι μεγαλύτερη

από του ανθρώπου. Το P είναι η πιθανότητα ένας άνθρωπος να ολοκληρώσει έναν γύρω και q η πιθανότητα να ολοκληρώσει έναν γύρο το λογισμικό, n είναι ο αριθμός των περιπτώσεων που το λογισμικό θα είναι πιο γρήγορο από τον άνθρωπο, m είναι ο αριθμός των γύρων και k είναι η δικλείδα των αριθμών των γύρων. Επομένως η αποτελεσματικότητα του CAPTCHA ορίζεται ως:

$$G = \sum_{i=k}^{m} \binom{m}{i} p^i (1-p)^{m-i} \cdot \left[1 - \sum_{i=k}^{m} \binom{m}{i} q^i (1-q)^{m-i} \right]^n$$

Εικόνα 2-22 Μετρώντας την αποτελεσματικότητα του CAPTCHA

1.10. ΕΦΑΡΜΟΓΕΣ

Όπως αναφέρθηκε και σε προηγούμενη ενότητα οι εφαρμογές που έχουν εντάξει στην φαρέτρα τους ένα CAPTCHA είναι όλο και περισσότερες. Συγκεκριμένα αυτές μπορούν να είναι:

- Online Δημοσκοπήσεις: Η δημοκρατία του διαδικτύου είναι μία σύγχρονη τάση δίνοντας την ευκαιρία στον επισκέπτη μίας σελίδας να εκφράσει την άποψή του πάνω σε ένα συγκεκριμένο θέμα. Οι ιδιοκτήτες των σελίδων με τον τρόπο αυτό επιθυμούν να καταγράψουν τις ειλικρινείς απόψεις των επισκεπτών τους και όχι αυτές των μηχανικών συστημάτων. Για παράδειγμα τον Νοέμβριο του 1999 η ιστοσελίδα slashdot.com διενήργησε μία Online δημοσκόπηση που αφορούσε στην ανάδειξη του καλύτερου Πανεπιστημίου που σχετίζεται με την πληροφορική. Αυτό που παρατηρήθηκε από τους υπευθύνους ήταν η εντατική συμμετοχή από συγκεκριμένες διευθύνσεις IP, ενώ στην ουσία οι υπεύθυνοι είχαν προσπαθήσει να περιορίσουν τη συμμετοχή μόνο σε μία ψήφο ανά IP. Η συγκεκριμένη δημοσκόπηση ουσιαστικά κατάντησε παρωδία, αφού στην ουσία τα συμμετέχοντα πανεπιστήμια του MIT και του Carnegie Mellon προσπάθησαν να διαμορφώσουν το αποτέλεσμα υπέρ τους μέσα από την παράκαμψη του συστήματος ασφαλείας της ιστοσελίδας. Για την ιστορία να αναφερθεί ότι το MIT απέσπασε 21.156 ψήφους, το Carnegie Mellon 21.032 και όλα τα υπόλοιπα λιγότερες από 1.000. Επομένως όσοι ενδιαφέρονται να αποσπάσουν

μία αξιόπιστη γνώμη θα πρέπει να θωρακίσουν την προσπάθειά τους μέσα από ορθές πρακτικές.

- Δωρεάν πάροχοι ηλεκτρονικού ταχυδρομείου: Διάφορες εταιρείες όπως η Yahoo! και η Microsoft δέχονται χιλιάδες επιθέσεις από bots καθημερινά δημιουργώντας πλαστούς εικονικούς λογαριασμούς. Η δημιουργία πλαστών λογαριασμών έχει ως αποτέλεσμα τον βομβαρδισμό των χρηστών με διαφημιστικά και βλαβερά μηνύματα τα οποία δημιουργούν αρνητική εικόνα για τον πάροχο και την αξιοπιστία του. Η Yahoo δίνει το δικαίωμα σε έναν χρήστη να δημιουργήσει λογαριασμό εφόσον μπορεί να επιλύσει γρίφους όπως ο παρακάτω.

Εικόνα 2-23: CAPTCHA από τη Yahoo

- Ρομπότ μηχανών αναζήτησης: Ορισμένες ιστοσελίδες δεν θέλουν να ευρετηριάζονται από μηχανές αναζήτησης. Υπάρχει μια ετικέτα HTML για να αποτρέπει τα bots μηχανών αναζήτησης από την ανάγνωση ιστοσελίδων, αλλά η ετικέτα δεν εγγυάται ότι τα bots δεν θα την διαβάσουν. Γι αυτό το λόγο οι ιστοσελίδες αυτές έχουν την ανάγκη εφαρμογής ενός μηχανισμού CAPTCHA ως δικλείδα ασφαλείας.
- Worms και spam: Τα CAPTCHA εφαρμόζονται εκτενώς στην περίπτωση αποτροπής δημιουργίας μηνυμάτων που δεν προέρχονται αό πραγματικό ηλεκτρονικό ταχυδρομείο.

1.11. ΠΛΕΟΝΕΚΤΗΜΑΤΑ / ΜΕΙΟΝΕΚΤΗΜΑΤΑ ΑΠΟ ΤΗ ΧΡΗΣΗ ΤΟΥΣ

Στην παρούσα ενότητα θα αναπτυχθούν τα βασικά πλεονεκτήματα και μειονεκτήματα του μηχανισμού. Συγκριμένα τα πλεονεκτήματα που μπορεί να αναγνωρίσει κανείς είναι τα παρακάτω:

- Προστασία κατά τη διαδικασία εγγραφής σε έναν ιστότοπο: Πολλές από τις δημοφιλείς ιστοσελίδες, όπως Yahoo, Facebook, Twitter, Hotmail, Gmail κτλ παρέχουν δωρεάν εγγραφή μέλους και γι αυτό το λόγο έχουν εντάξει στη φαρέτρα των μηχανισμών ασφαλείας που διαθέτουν το CAPTCHA. Γι αυτό

το λόγο είναι απαραίτητη προϋπόθεση να εξασφαλιστεί ότι κάθε νέα εγγραφή στον ιστότοπο πραγματοποιείται από κάποιον ενδιαφερόμενο και όχι από ένα λογισμικό.

- Προστασία από spam: Το φαινόμενο των spam μηνυμάτων είναι διαδεδομένο και σίγουρα πολύ ενοχλητικό. Προς το παρόν χιλιάδες spammers μπορούν και διοχετεύουν στο διαδίκτυο εκατομμύρια μηνύματα με ελάχιστο κόπο. Τα spam μηνύματα σίγουρα δεν ενδιαφέρουν κανέναν και το κυριότερο μπορούν να δημιουργήσουν αναστάτωση στην ηλεκτρονική αλληλογραφία του χρήστη.
- Ασφαλέστερη εμπειρία αγορών: Η διασφάλιση των παραγγελιών από την πλευρά των ιδιοκτητών των online καταστημάτων είναι πολύ σημαντική. Τα μηνύματα που δέχονται καθημερινώς για παραγγελίες θα πρέπει να αφορούν σε ανθρώπους και όχι σε κακόβουλο λογισμικό. Σε αντίθετη περίπτωση ιστοσελίδες εμπορικού περιεχομένου που δεν έχουν διασφαλίσει στο ελάχιστο τη φύση των πελατών τους είναι καταδικασμένες να καταναλώνουν πολύτιμο χρόνο στην εξακρίβωση των στοιχείων τους.
- Προστασία του ηλεκτρονικού ταχυδρομείου: Οι πιθανές επιθέσεις από hackers μπορούν να θέσουν στο στόχαστρο τα ηλεκτρονικά ταχυδρομεία των χρηστών Για αυτό το λόγο έχει τεθεί ως δικλίδα ασφαλείας το CAPTCHA προκειμένου να εξακριβωθεί ότι αυτός που προσπαθεί να μπει στο λογαριασμό είναι άτομο και όχι λογισμικό.

Μειονεκτήματα

Στην προηγούμενη ενότητα αναφέρθηκαν τα πλεονεκτήματα του συγκεκριμένου μηχανισμού, αλλά σίγουρα όπως προαναφέρθηκε δεν αποτελεί πανάκεια ούτε τον απόλυτο μηχανισμό ασφαλείας. Ουσιαστικά τα μειονεκτήματα του μπορούν να συνοψιστούν στα εξής τρία σημεία.

- Τα CAPTCHAs μπορούν να σπάσουν. Όπως προαναφέρθηκε με τον έναν ή τον άλλο τρόπο ο μηχανισμός έχει κενά που ένα καλά μελετημένο λογισμικό μπορεί να εντοπίσει και να χτυπήσει. Όμως στη διαδικασία καταπολέμησης του CAPTCHA δεν παίζει ρόλο μόνο η μηχανή, αλλά και ο άνθρωπος. Χιλιάδες εργαζόμενοι καθισμένοι μπροστά στον υπολογιστή τους αποτελούν την ομάδα του deCAPTCHA, οι οποίοι προσπαθούν να σπάσουν περιπτώσεις CAPTCHA και να τροφοδοτήσουν με υλικό εκμάθησης το λογισμικό. Βέβαια

ακόμα και αν δεν υπήρχε η ανθρώπινη συμβολή η προσπάθεια έκθεσης των μηχανισμών αυτών γίνεται και από αλγορίθμους που προσπαθούν να κατανοήσουν τη συμπεριφορά τους. Οι αλγόριθμοι αυτοί αναπτύσσονται από ομάδες τεχνικών και πωλούνται έναντι αδρής αμοιβής. Δύσκολα θα βρεθεί κάποιος αποτελεσματικός αλγόριθμος ελεύθερος στο διαδίκτυο. Πάντως υπηρεσίες όπως αυτές που προσφέρουν οι ομάδες deCAPTCHA γίνονται ολοένα και πιο διαδεδομένες.

- Τα CAPTCHAs μπορούν να αποτρέψουν τους πραγματικούς χρήστες. Πολλοί από τους χρήστες του διαδικτύου θεωρούν δύσχρηστο το CAPTCHA και πολλές φορές δεν μπαίνουν στον κόπο να συνεχίσουν τη δραστηριότητά τους εξαιτίας του. Μάλιστα πολλές φορές δεν μπαίνουν στη διαδικασία να ξαναεπισκεφθούν την ιστοσελίδα θεωρώντας την και αυτή δύσχρηστη όπως το CAPTCHA. Ταυτόχρονα θα πρέπει να ληφθεί υπόψην και το πρόβλημα όρασης που μπορεί να αντιμετωπίζουν ορισμένα άτομα. Το CAPTCHA βασίζεται στην καλή όραση και οποιοσδήποτε δεν διαθέτει δεν μπορεί να ανταποκριθεί.
- Το CAPTCHA είναι απαιτητικό. Όλες οι ιστοσελίδες που χρησιμοποιούν CAPTCHAs θα πρέπει να αφιερώσουν αρκετές από τις δυνατότητές τους στην φιλοξενία του μηχανισμού. Για τη φιλοξενία των μηχανισμών απαιτούνται πολλοί πόροι και servers με μεγάλες δυνατότητες. Ιστοσελίδες με περιορισμένες δυνατότητες ίσως δεν μπορούν να υποστηρίξουν τέτοιου είδους απαιτήσεις, γεγονός που μάλλον αποτελεί μειονέκτημά τους.

1.12. ΣΥΝΟΨΗ ΚΕΦΑΛΑΙΟΥ

Το έντονο ενδιαφέρον τόσο του επιχειρηματικού κόσμου όσο και της επιστημονικής κοινότητας για την ανάπτυξη νέων τεχνικών που αφορούν την επίλυση των μεθόδων CAPTCHA. Ο στόχος των πρώτων είναι να αυξήσουν την επιρροή τους στη δεξαμενή των καταναλωτών. Από την άλλη στόχος της επιστημονικής κοινότητας είναι να αυξήσει τόσο την ασφάλεια των μεθόδων που αναπτύσσει όσο και την ευχρηστία. Οι λύσεις που προσφέρονται είναι πολλές και η επιλογή εξαρτάται αποκλειστικά από την εμπιστοσύνη που δείχνε ο ιδιοκτήτης μιας ιστοσελίδας σε κάθε μία από αυτές. Σκοπός της παρούσας εργασίας δεν είναι να ταχτεί υπέρ της μιας ή της

άλλης μεθόδου, αλλά να παρουσιάσει τις υπάρχουσες και να προτείνει ένα δικό της μοντέλο επίλυσης ενός CAPTCHA.

Η κάλυψη των δύο παραπάνω στόχων θα βελτιώσει την αντιλάμβανόμενη αξία της παρεχόμενης υπηρεσίας και θα αυξήσει τα κέρδη για τις εταιρείες που παρέχουν παρόμοιες υπηρεσίες. Με τον τρόπο αυτό θα αυξηθεί ο τζίρος στο συγκεκριμένο οικονομικό πεδίο και θα στραφεί το ενδιαφέρον στη βελτιστοποίηση της παρεχόμενης υπηρεσίας.

ΠΗΓΕΣ ΚΕΦΑΛΑΙΟΥ

- Asirra - http://research.microsoft.com/en-us/um/redmond/projects/asirra/

- Bigham, J.P. and A.C. Cavender, 2009. Evaluating existing audio CAPTCHAs and an interface optimized for non-visual use. Proceedings of the 27th International Conference on Human Factors in Computing Systems, April 4-9, 2009, Boston, MA., USA., pp: 1829-1838.

- Chew Monica and. Tygar J. D, UC Berkeley (2004) Image Recognition CAPTCHAs In Proceedings of the 7th International Information Security Conference (ISC 2004), Springer, September 2004, pp. 268-279

- Cryptographp - http://www.CAPTCHA.fr/

- Elson, Jeremy Douceur, John R. Howell, Jon and Saul. Jared (2007) Asirra: a CAPTCHA that exploits interest-aligned manual image categorization. In ACM Conference on Computer and Communications Security, pages 366–374. ACM, 2007.

- **FreeCap -** http://www.freecap.ru/eng/

- Gupta, A., A. Jain, A. Raj and A. Jain, 2009. Sequenced tagged CAPTCHA: Generation and its analysis. Proceedings of the IEEE International Advance Computing Conference, March 6-7, 2009, Patiala, Punjab, India, pp: 1286-1291

- Chu, B.B., J. Yan, Q. Li, C. Yang and J. Liu *et al*., 2010. Attacks and design of image recognition CAPTCHAs. Proceedings of the 17th ACM Conference on Computer and Communications Security, October 4-8, 2010, Chicago, IL., USA., pp: 187-200.

- Jeff Yan and Ahmad Salah El Ahmad. 2008. Usability of CAPTCHAs or usability issues in CAPTCHA design. In *Proceedings of the 4th symposium on Usable privacy and security* (SOUPS '08). ACM, New York, NY, USA, 44-52

- Milde B. (2010) On the security of re CAPTCHA Bachelor-Thesis from Darmstadt

- Motoyama, Marti, Levchenko, Kirill, Kanich, Chris McCoy, Damon Voelker, Geoffrey M. and Savage, Stefan. (2010). Re: CAPTCHAs: understanding

CAPTCHA-solving services in an economic context. In Proceedings of the 19th USENIX conference on Security (USENIX Security'10). USENIX Association, Berkeley, CA, USA, 28-28.

- Nielsen Jakob. (2003) Usability 101: Introduction to Usability. Διαθέσιμο στο http://www.useit.com/alertbox/20030825.html.

- NuCAPTCHA - http://www.nuCAPTCHA.com/

- Rolko Juraj (2010) 3D CAPTCHA: CAPTCHA based on spatial perspective and human imagination Διαθέσιμο στο http://www.3dCAPTCHA.net/documents/3D_CAPTCHA.pdf

- Securimage - http://www.phpCAPTCHA.org/

- Symantec. A captcha-solving service. http://www.symantec.com/connect/blogs/captcha-solving-service

- TextCAPTCHA - http://textCAPTCHA.com/

- Webspamprotect - http://webspamprotect.com/

- wp-CAPTCHA - http://wordpress.org/plugins/wp-CAPTCHA-free/

.

– ΔΗΜΙΟΥΡΓΙΑ ΙΣΤΟΤΟΠΟΥ

Για την δημιουργία του δικτυακού τόπου θα χρησιμοποιήσουμε ένα σύστημα διαχείρισης περιεχομένου (CMS: **C**ontent **M**anagement **S**ystem) το Joomla.

Ένα **Σ**ύστημα **Δ**ιαχείρισης **Π**εριεχομένου, περιέχει όλα εκείνα τα εργαλεία, για την δημιουργία ιστοτόπων εύκολα και γρήγορα. Μπορούν να χρησιμοποιηθούν για την κατασκευή οποιουδήποτε είδους Η κατασκευή μπορεί να περιλαμβάνει όλων των ειδών των δικτυακών τόπων όπως προσωπικές σελίδες, ημερολόγια ιστού (blogs), εμπορικά δικτυακά καταστήματα (eshops) ακόμη και κοινωνικές σελίδες δικτύωσης.

Το περιεχόμενο των ιστοτόπων είναι δυναμικό, αποθηκεύεται σε βάσεις δεδομένων και μπορεί να ανανεώνεται γρήγορα και εύκολα χωρίς να απαιτούνται ιδιαίτερες γνώσεις προγραμματισμού.

1.13. ΣΥΣΤΗΜΑΤΑ ΔΙΑΧΕΙΡΙΣΗΣ ΠΕΡΙΕΧΟΜΕΝΟΥ

Με την ανάπτυξη του διαδικτύου υπήρξε ανάγκη δημιουργίας αποδοτικότερων εργαλείων δημιουργίας ιστοσελίδων. Η φράση “catch-all” επινοήθηκε για να περιγράψει το ευρύ φάσμα των συστημάτων που επέτρεπαν στους χρήστες την δημιουργία, επεξεργασία και διαχείριση του περιεχομένου μιας ιστοσελίδας.

Στις αρχές του 1990, υπήρξε η πρώτη δυνατότητα ενημέρωσης του διαδικτυακού περιεχομένου με προϊόντα που είχαν αναπτυχθεί από την Microsoft και την Lotus, ενώ το πρώτο υποτυπώδες σύστημα διαχείρισης περιεχομένου ήταν το StoryServer από την Vignette γύρω στο 1996. Τα χρόνια που ακολούθησαν παρουσιάστηκαν αρκετές προτάσεις προετοιμάζοντας για την ανάπτυξη που θα είχαν σήμερα τα CMS.

Παρ’ όλα αυτά μεταξύ του 2000 και 2005 υπήρξε ένα μαζικό κύμα συγχωνεύσεων και εξαγορών μεταξύ των εταιρειών που δραστηριοποιούνταν στο συγκεκριμένο χώρο. Τα συστήματα που είχαν δημιουργηθεί εγκαταλείφτηκαν με αποτέλεσμα οι χρήστες να μείνουν χωρίς υποστήριξη.

Η αγορά όμως γρήγορα επανέκαμψε και το 2007 υπήρχαν ουσιαστικά τρεις τύποι συστημάτων:

- **Αποσύνδεσης**: Εξυπηρετούσαν την σχεδίαση του ιστοτόπου σε τοπικό επίπεδο και στη συνέχεια γίνονταν μεταφόρτωση του περιεχομένου στο διαδίκτυο. Απαιτούσαν την εγκατάσταση κάποιας εφαρμογής τοπικά.

- **Απ' ευθείας σύνδεσης**: Αυτά τα συστήματα δεν χρειαζόντουσαν εγκατάσταση κάποιου προγράμματος, αφού το πρόγραμμα διαχείρισης ήταν από μόνο του ένας δικτυακός τόπος που μπορούσες να συνδεθείς με όνομα χρήστη και συνθηματικά.

- **Υβριδικά συστήματα**: Υποστήριζαν και τις δύο λειτουργίες (on-line, off-line).

Σήμερα τα CMS έχουν γίνει εξαιρετικά πολύπλοκα, επιτρέποντας στους χρήστες να κάνουν σχεδόν τα πάντα, χωρίς την συγγραφή ούτε μιας γραμμής κώδικα. Οι νέες εξελίξεις έχουν ενσωματώσει πρακτικές μάρκετινγκ όπως συστήματα αποστολής μαζικών μηνυμάτων μέσω ενσωματωμένου ηλεκτρονικού ταχυδρομείου, παρακολούθηση στατιστικών στοιχείων, δημιουργία δημοσκοπήσεων και μια σειρά άλλων δυνατοτήτων.

Ένα Σύστημα Διαχείρισης Περιεχομένου χωρίζει το σχεδιασμό και τη διαχείριση της ιστοσελίδας από την προβολή του περιεχομένου. Εάν προσπαθήσουμε να δημιουργήσουμε ένα στατικό δικτυακό τόπο θα ανακαλύψουμε την δυσκολία ενημέρωσής του με νέες σελίδες. Ακόμα πιο επίπονη διαδικασία είναι η προσθήκη διαφορετικών μενού σε αυτές τις νέες σελίδες. Ένα Σύστημα Διαχείρισης Περιεχομένου αναλαμβάνει την αυτοματοποίηση αυτών των διαδικασιών και αφήνει τον δημιουργό να επικεντρωθεί αυστηρά στο περιεχόμενο.

Σύμφωνα με την Wikipedia:

> *" Τα Συστήματα Διαχείρισης Περιεχομένου (ΣΔΠ, Content Management Systems, CMS) είναι διαδικτυακές εφαρμογές που επιτρέπουν την online τροποποίηση του περιεχομένου ενός δικτυακού τόπου. Οι διαχειριστές μέσω του διαδικτύου ενημερώνουν το περιεχόμενο στο ΣΔΠ, το οποίο είναι εγκατεστημένο σ' ένα διακομιστή. Οι αλλαγές αυτές γίνονται αυτόματα διαθέσιμες πάλι μέσω του διαδικτύου, σε όλους τους επισκέπτες και χρήστες του δικτυακού τόπου"*

Τα εργαλεία αυτού του συστήματος χρησιμοποιούνται για να παρέχουν:

- **Ασφάλεια στην πρόσβαση**: Με την δημιουργία χρηστών και δικαιωμάτων μπορούμε να καθορίσουμε το επίπεδο πρόσβασης. Έτσι μπορούμε να αναθέσουμε συγκεκριμένες αρμοδιότητες σε ένα άτομο ή να απαγορέψουμε ανάλογα κάποιες. Για παράδειγμα μπορούμε να δημιουργήσουμε ένα χρήστη

ο οποίος θα μπορεί να τροποποιήσει ένα άρθρο, αλλά δεν θα μπορεί να το σβήσει.

- **Εύκολη διαχείριση των δεδομένων**: Εφόσον όλα τα δεδομένα αποθηκεύονται σε μια βάση δεδομένων, η πρόσβαση και η διαχείριση αυτών γίνεται αρκετά εύκολα. Η επικαιροποίηση των δεδομένων συνιστά μια απλή διαδικασία
- **Γεννήτρια αναφορών και στατιστικών**: Η οργάνωση των δεδομένων, μας δίνει την δυνατότητα παροχής στατιστικών στοιχείων. Επίσης η εκτύπωση αναφορών γίνεται εύκολα μέσω ερωτημάτων στην βάση.
- **Υποστήριξη για πολυμέσα**: Υπάρχει η δυνατότητα διαχείρισης οποιουδήποτε αρχείου όπως απλά έγγραφα, ταινίες, εικόνες.

Τα συστήματα διαχείρισης περιεχομένου, συνήθως αποτελούνται από δύο μέρη:

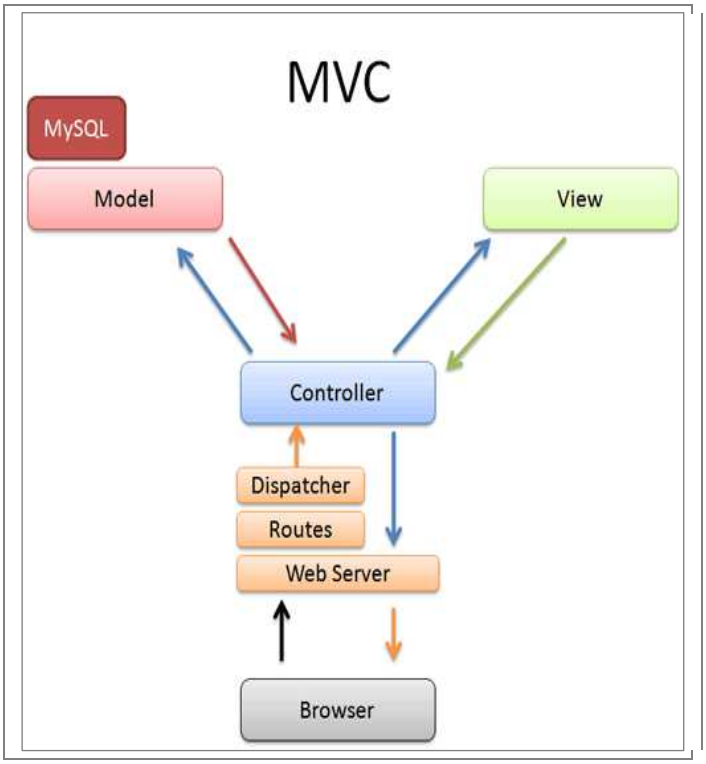
- **back-end**: Αποτελεί το σύστημα διαχείρισης του Joomla, που παρέχει όλα εκείνα τα εργαλεία για την δημιουργία του ιστοτόπου.
- **front-end**: Είναι το τμήμα που βλέπει ο επισκέπτης, δηλαδή ο ίδιος ο ιστότοπος. Ουσιαστικά το front-end είναι το αποτέλεσμα των ενεργειών που γίνονται στο back-end.

Οι συνηθέστερες ενέργειες που γίνονται στο back-end είναι:

- **Ενεργοποίηση συστήματος εισόδου**: Χρησιμοποιείται για την δημιουργία μιας οθόνης εισόδου για την διαπίστευση του χρήστη κατά την επίσκεψή του στον δικτυακό τόπο. Επίσης δίνει την δυνατότητα εγγραφής ενός χρήστη. Ο εγγεγραμμένος χρήστης μπορεί να χρήζει ιδιαίτερων δικαιωμάτων, όπως για παράδειγμα η επίσκεψη σε σελίδες που δεν μπορεί να δει ο απλός επισκέπτης του ιστοτόπου.
- **Ενεργοποίηση συστήματος δημοφιλέστερων ή νεότερων άρθρων**: Παρέχει την δυνατότητα εμφάνισης των δημοφιλέστερων στοιχείων σε έναν δικτυακό τόπο (άρθρων, εικόνων, ταινιών κλπ).
- **Δημιουργία ψηφοφοριών**: Παρέχει την δυνατότητα δημοσκοπήσεων.
- **Παροχή πεδίου αναζήτησης**: Παρέχει την δυνατότητα αναζήτησης των περιεχομένων του ιστότοπου.
- **Πολύγλωσση υποστήριξη**: Παρέχει την δυνατότητα δημιουργίας ιστοσελίδων σε πολλές γλώσσες.

1.14. ΧΑΡΑΚΤΗΡΙΣΤΙΚΑ JOOMLA

Το Joomla σε επίπεδο εφαρμογής ακολουθεί το σχέδιο Μοντέλο – Προβολή – Έλεγχος (**MVC**, **M**odel-**V**iew-**C**ontroller).



Εικόνα 3-24 - Μοντέλο MVC

(πηγή: Intermediate Rails: Understanding Models, Views and Controllers, 2011)

Το μοντέλο MVC αποτελείται όπως φαίνεται από την εικόνα 1 από τις εξής ενότητες:

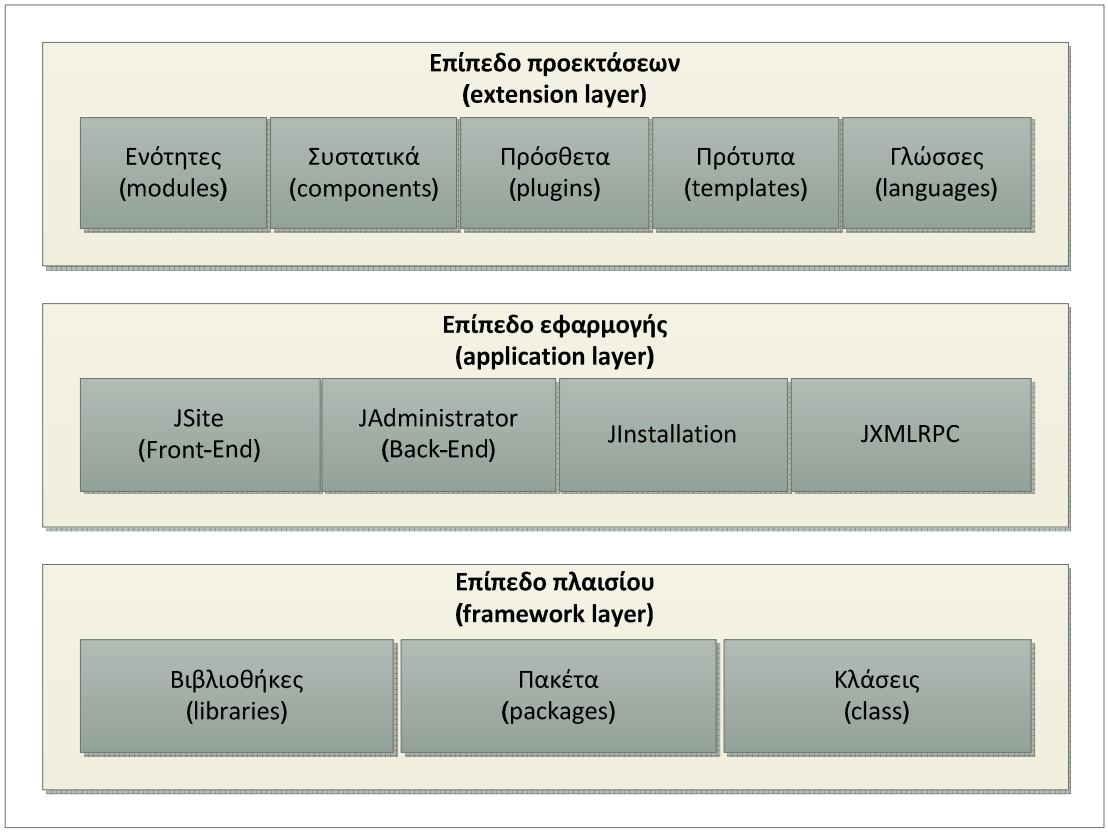
- **Model:** Αναπαριστά τα δεδομένα της εφαρμογής
- **View:** Παράγει μια παρουσίαση των δεδομένων του μοντέλου
- **Controller:** Διαχειρίζεται και κατευθύνει τα αιτήματα των χρηστών

Η διαδικασία που πραγματοποιείται όταν δημιουργείται ένα αίτημα, δηλαδή η ανάκτηση μιας σελίδας συνοπτικά είναι:

1. Ο φυλλομετρητής στέλνει ένα αίτημα για την ανάκτηση μιας σελίδας.
2. Ο εξυπηρετητής (Web Server) λαμβάνει το αίτημα και το δρομολογεί.
3. Στη συνέχεια ο controller δίνει τις κατάλληλες εντολές, για να ανακτηθεί η αντίστοιχη πληροφορία, μέσω του model.

4. Το model αναλαμβάνει να προσπελάσει την βάση για την ανάκτηση της πληροφορίας.

Τέλος το view είναι υπεύθυνο για την παρουσίαση της ανακτημένης πληροφορίας στην κατάλληλη μορφή. Το Joomla είναι ένα σύστημα αρχιτεκτονικής τριών βαθμίδων (3 tier system). Αποτελείται τρία επίπεδα: Το επίπεδο προεκτάσεων, το επίπεδο εφαρμογής και το επίπεδο πλαισίου [Εικόνα 3-25 – Αρχιτεκτονική Joomla 1.5].



Εικόνα 3-25 – Αρχιτεκτονική Joomla 1.5

1.14.1. Επίπεδο προεκτάσεων

Το επίπεδο προεκτάσεων είναι υπεύθυνο για την επέκταση των δυνατοτήτων του Joomla και των εφαρμογών του. Αποτελείται από τα ενθέματα (modules), τα συστατικά (components), τα πρόσθετα (plugins), τα πρότυπα (templates) και τις γλώσσες (languages).

- **Συστατικά (components)**: Τα συστατικά είναι εφαρμογές που επεκτείνουν το πλαίσιο του Joomla. Υπάρχει ένα βασικό σετ συστατικών, ενσωματωμένο στον πυρήνα του Joomla για την εκτέλεση διάφορων λειτουργιών.
- **Ενθέματα** (**modules**): Τα ενθέματα επιτρέπουν την εκτέλεση εργασιών σε καθορισμένες περιοχές της ιστοσελίδας. Εμφανίζονται σε διάφορες θέσεις και ενεργοποιούνται ή απενεργοποιούνται ανάλογα με την σελίδα που βρίσκεται ο χρήστης. Οι θέσεις των ενθεμάτων εξαρτώνται από το εκάστοτε πρότυπο που χρησιμοποιείται. Μπορούν να τοποθετηθούν παραπάνω από ένα ένθεμα στην ίδια θέση. Κάθε ένα τοποθετείται οριζοντίως κάτω από το προηγούμενο.
- **Πρόσθετα (plugins):** Τα πρόσθετα αποτελούν ένα είδος επέκτασης των δυνατοτήτων του Joomla. Τα πρόσθετα παρέχουν ρουτίνες οι οποίες ενεργοποιούνται βάσει κάποιου γεγονότος. Όταν ένα συγκεκριμένο γεγονός συμβεί όλες οι ρουτίνες τους πρόσθετου εκτελούνται στη σειρά. Με αυτό τον τρόπο προστίθεται λειτουργικότητα στο πλαίσιο του Joomla.
- **Πρότυπα (templates):** Το Joomla έρχεται με μία σειρά βασικών προτύπων για την εμφάνιση του ιστότοπου. Αυτά τα πρότυποα είναι τα: Rhuk_Milkyway, Beez και JA_Purity. Η αλλαγή των προτύπων μπορεί να γίνει δυναμικά, βοηθώντας έτσι τον διαχειριστή να επιλέξει για την καλύτερη αισθητική του δικτυακού τόπου.

1.14.2. Επίπεδο εφαρμογής

Η μεσαία βαθμίδα (application layer), αποτελείται από εφαρμογές που επεκτείνουν την βασική κλάση JApplication. Το πλαίσιο Joomla αποτελείται από τέσσερις βασικές εφαρμογές:

- **JInstallation**: Η εφαρμογή JInstallation είναι υπεύθυνη για την εγκατάσταση του Joomla σε έναν εξυπηρετητή WEB. Αποτελείται από έναν οδηγό που καθοδηγεί τον χρήστη στις διάφορες επιλογές. Μόλις τελειώσει η διαδικασία της εγκατάστασης η εφαρμογή διαγράφεται για λόγους ασφάλειας.
- **JSite**: Η εφαρμογή JSite είναι υπεύθυνη για την διεπιφάνεια του διαδικτυακού χώρου. Μέσω των διάφορων λειτουργιών που υποστηρίζει βοηθάει στην αλληλεπίδραση των χρηστών με το back-end του συστήματος (π.χ. φόρμα εισαγωγής στοιχείων).

- **JAdministrator**: Η εφαρμογή αυτή αποτελεί το back-end του συστήματος. Μέσα από ένα γραφικό περιβάλλον γίνεται η διαχείριση του ιστοτόπου και η επεξεργασία των στοιχείων που συλλέγονται από το front-end.
- **JXMLRPC**: Η εφαρμογή αυτή είναι υπεύθυνη να παρέχει τις διαδικασίες, οι οποίες εκτελούν λειτουργίες απομακρυσμένης διαχείρισης, σε έναν δικτυακό τόπο που έχει σχεδιαστεί σε Joomla.

1.14.3. Επίπεδο πλαισίου

Είναι το κατώτερο επίπεδο στην αρχιτεκτονική Joomla και αποτελείται από βιβλιοθήκες, πακέτα και κλάσεις που βοηθούν στην ανάπτυξη κώδικα για την δημιουργία πρόσθετων ή άλλων χαρακτηριστικών για την επέκταση του Joomla. Οι κλάσεις και τα πακέτα ανάλογα με τον σκοπό που επιτελούν χωρίζονται σε διαφορετικές υποκατηγορίες όπως[1]: Εφαρμογή, Σύστημα αρχείων, Εργαλεία, Χρήστης, Περιβάλλον, Html, Προσωρινός χώρος αποθήκευσης, Βάση δεδομένων, Έγγραφο, Συμβάν, Λάθος, Γλώσσα, Πελάτης, Βασικό, Πρόσθετο, Μητρώο, Πρότυπο και ενότητα.

Το πλαίσιο Joomla είναι το πιο σημαντικό μέρος την αρχιτεκτονικής Joomla. Βασίζεται σε αντικειμενοστραφή λογική και αυτό κάνει την διαχείριση και την επεκτασιμότητα εύκολη.

1.14.4. Περιεχόμενο (Content)

Το Joomla υποστηρίζει την ιεραρχική δομή στην οργάνωση και εμφάνιση του περιεχομένου. Στο ανώτερο επίπεδο βρίσκονται οι ενότητες. Οι ενότητες δεν περιέχουν άρθρα, αλλά κατηγορίες που βρίσκονται σε επόμενο προς τα κάτω επίπεδο. Οι κατηγορίες περιέχουν άρθρα. Κάθε κατηγορία μπορεί να ανήκει σε μόνο μία ενότητα και κάθε άρθρο σε μόνο μία κατηγορία. Επίσης υπάρχει η δυνατότητα ένα άρθρο να μην ανήκει σε καμία κατηγορία.

Η παραπάνω δομή έχει δύο βασικά πλεονεκτήματα: το πρώτο είναι λόγω των εσωτερικών συνδέσμων που δημιουργούνται ο ιστότοπος είναι φιλικός προς τις μηχανές αναζήτησης και το δεύτερο ότι επιτρέπεται η παρουσίαση του περιεχομένου με πολλούς τρόπους:

[1] Μια πλήρη λίστα του Joomla API βρίσκεται στην διεύθυνση: http://api.joomla.org/li_Joomla-Framework.html

- **Κατάλογος αρχειοθετημένων άρθρων**: Εμφανίζει μια λίστα με το τίτλο των αρχειοθετημένων άρθρων κατά χρονολογική σειρά (αύξουσα ή φθίνουσα).
- **Σελιδοποίηση άρθρου**: Εμφανίζει ένα συγκεκριμένο άρθρο.
- **Σελιδοποίηση υποβολής άρθρου**: Επιτρέπει στους χρήστες που είναι αρθογράφοι να υποβάλλουν κάποιο άρθρο.
- **Σελιδοποίηση κατηγορίας ως ιστολόγιο**: Εμφανίζει τα άρθρα της κατηγορίας με χρονολογική σειρά (αύξουσα ή φθίνουσα).
- **Βασική σελιδοποίηση κατηγορίας**: Εμφανίζει τα άρθρα μιας κατηγορίας με τη μορφή περιεχομένων.
- **Σελιδοποίηση πρωτοσέλιδου ως ιστολόγιο**: Εμφανίζει τα άρθρα της πρώτης σελίδας με χρονολογική σειρά (αύξουσα ή φθίνουσα).
- **Σελιδοποίηση ενότητας ως ιστολόγιο**: Εμφανίζει τα άρθρα των κατηγοριών μιας ενότητας με χρονολογική σειρά (αύξουσα ή φθίνουσα).
- **Σελιδοποίηση Ενότητας**: Εμφανίζει μια λίστα με τις κατηγορίες μιας ενότητας.

1.14.5. Μενού (Menus)

Το Joomla προσφέρει την δυνατότητα πλοήγησης των χρηστών στον δικτυακό τόπο μέσω ενός συστήματος μενού. Τα προκαθορισμένα μενού του Joomla είναι τέσσερα:

- **Top menu**: Το μενού αυτό είναι τοποθετημένο στο πάνω μέρος της οθόνης σε οριζόντια διάταξη.
- **Main menu**: Τοποθετείται εξ' ορισμού στο αριστερό μέρος της οθόνης σε κάθετη διάταξη.
- **Other menu**: Τοποθετείται στο κάτω αριστερό μέρος της οθόνης και περιλαμβάνει συνδέσμους σε εξωτερικούς ιστότοπους.
- **User menu**: Το μενού αυτό εμφανίζεται όταν συνδέεται ένας χρήστης στον δικτυακό τόπο και περιλαμβάνει πάντα μια επιλογή αποσύνδεσης.

1.14.6. Χρήστες (Users)

Χρήστης είναι οποιοσδήποτε επισκέπτεται τον δικτυακό τόπο. Ένας χρήστης μπορεί να είναι επισκέπτης ή πιστοποιημένος. Ο πιστοποιημένος χρήστης πρέπει να έχει λογαριασμό και να συνδέεται με τα διαπιστευτήριά του στον ιστότοπο. Οι χρήστες ανήκουν σε μία από τις επτά προκαθορισμένες ομάδες όπως: Υπερδιαχειριστής, Διαχειριστής, Συντονιστής για την δημόσια διαχείριση και Μέλος, Αρθρογράφος, Συντάκτης, Εκδότης για την διαχείριση του περιεχομένου του δικτυακού τόπου. Για τους εγγεγραμμένους χρήστες υπάρχει η δυνατότητα αλλαγής των στοιχείων τους, μέσω ενός μενού επιλογών. Επίσης μπορούν να δημοσιεύσουν και μια συζήτηση στον φόρουμ του ιστοτόπου.

Οι επισκέπτες του δικτυακού τόπου έχουν την δυνατότητα πρόσβασης σε όλες τις σελίδες. Σε αντίθεση με τους εγγεγραμμένους χρήστες, δεν μπορούν όμως να ανεβάσουν κάποιο άρθρο, ενώ στην ενότητα του φόρουμ μπορούν μόνο να δουν τις συζητήσεις και όχι να δημιουργήσουν.

1.15. ΑΠΑΙΤΗΣΕΙΣ ΣΥΣΤΗΜΑΤΟΣ

Οι απαιτήσεις ενός συστήματος για την δημιουργία ενός δικτυακού τόπου σε Joomla, είναι:

- **Web Server Apache**: Για την προσπέλαση του δικτυακού τόπου είναι απαραίτητη η εγκατάσταση ενός εξυπηρετητή WEB όπως είναι ο Apache.
- **PHP**: Για την λειτουργία του Joomla είναι απαραίτητη η συγκεκριμένη γλώσσα προγραμματισμού, αφού το Joomla έχει δημιουργηθεί σε PHP.
- **mySQL**: Σε mysql θα φιλοξενηθεί η βάση του συστήματος του Joomla.
- **Joomla**: Η τελευταία σταθερή έκδοση του Joomla η οποία είναι η 3.0.3

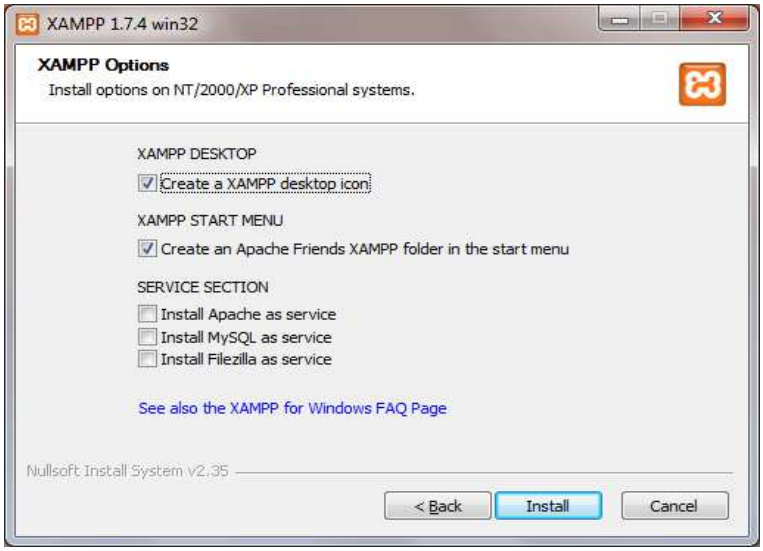
1.16. ΕΓΚΑΤΑΣΤΑΣΗ JOOMLA

Εφόσον πληρούνται όλες οι απαραίτητες προϋποθέσεις η εγκατάσταση του Joomla είναι μια αρκετά εύκολη διαδικασία. Επίσης η εξασφάλιση των προαπαιτούμενων, για την εγκατάσταση μπορεί να γίνει με την χρήση του πακέτου XAMPP.

Το XAMPP είναι ένα πακέτο που περιλαμβάνει όλες τις ανωτέρω εφαρμογές οι οποίες είναι απαραίτητες για την εγκατάσταση και λειτουργία του Joomla. Η έκδοση

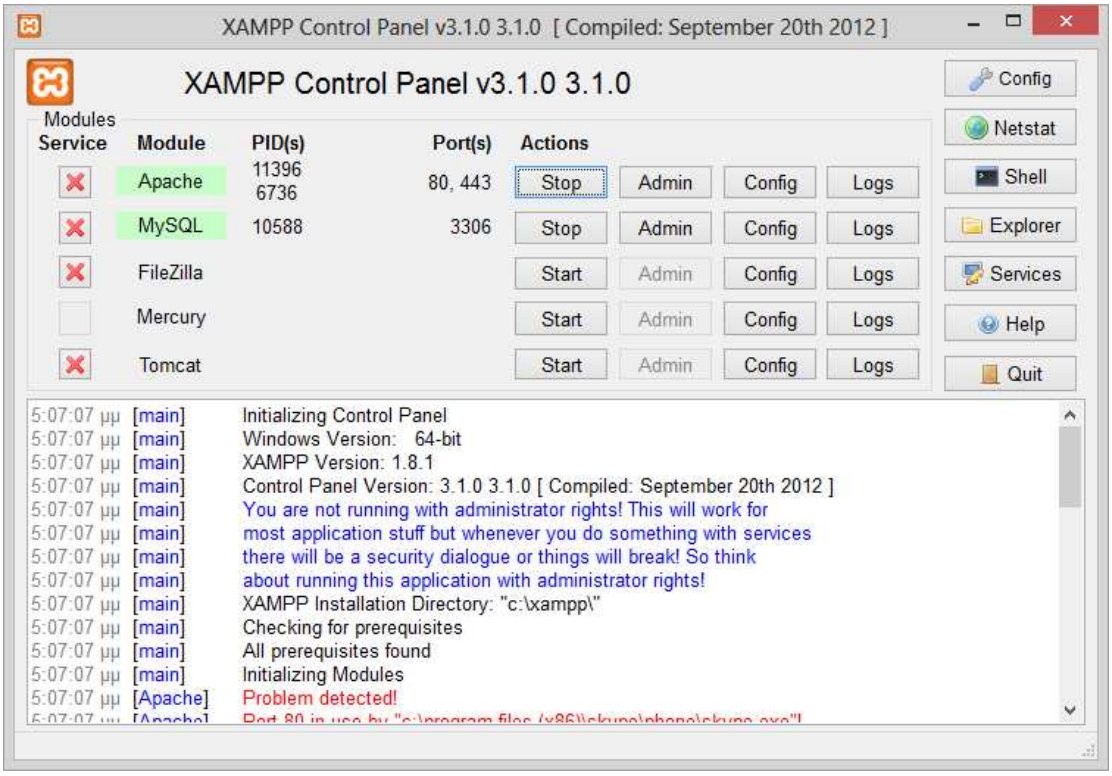
που θα χρησιμοποιήσουμε είναι για περιβάλλον Windows και η εγκατάσταση θα γίνει σε τοπικό επίπεδο.

Αφού κατεβάσουμε το XAMPP εκτελούμε το αρχείο εγκατάστασης και ακολουθούμε τα βήματα.



Εικόνα 3-26-Xampp Options

Η διαδικασία είναι αρκετά απλή και μόλις εγκατάσταση τελειώσει, εκτελούμε τον Πίνακα Ελέγχου του XAMPP (XAMPP Control Application) Στον πίνακα ελέγχου ενεργοποιούμε τις υπηρεσίες που μας ενδιαφέρουν δηλαδή: Apache και MySQL όπως βλέπουμε στην εικόνα 5.



Εικόνα 3-27 Πίνακας ελέγχου XAMPP

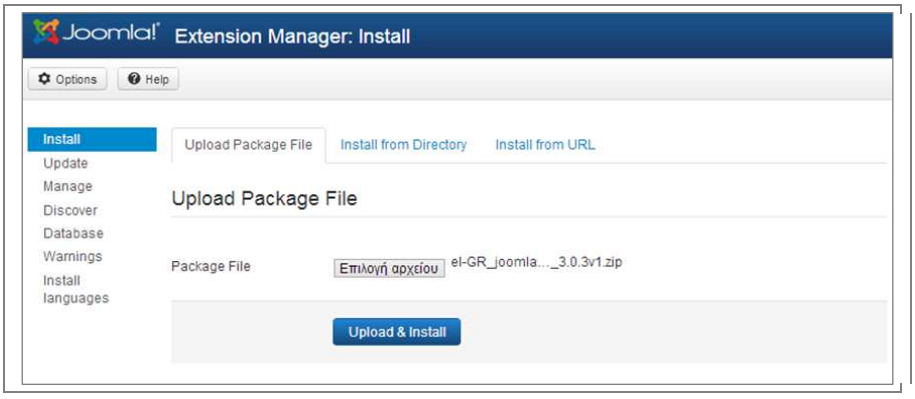
Μόλις οι υπηρεσίες μας εκκινήσουν, μπορούμε στην διεύθυνση http://127.0.0.1/xampp/index.php, να δούμε την ακόλουθη εικόνα, η οποία μας δείχνει ότι όλα πήγαν καλά με την εγκατάσταση.

Εικόνα 3-28 Οθόνη καλωσορίσματος του XAMPP

1.17. ΥΛΟΠΟΙΗΣΗ ΜΕΣΩ ΤΗΣ ΔΙΑΧΕΙΡΙΣΗΣ (frontend)

1.17.1. Εγκατάσταση ελληνικής γλώσσας

Το πρώτο βήμα είναι η εγκατάσταση της ελληνικής γλώσσας. Αυτή γίνεται αφού πρώτα κατεβάσουμε το αντίστοιχο αρχείο της ελληνικής διανομής. Η εγκατάσταση γίνεται μέσα από το μενού "Extension Manager → Install → Upload Package File." Επιλέγουμε το αντίστοιχο αρχείο και πατάμε το "Upload & Install".



Εικόνα 3-29 Εγκατάσταση ελληνικής γλώσσας

1.17.2. Πρόσθετα

Η εγκατάσταση προσθέτων όπως είδαμε επεκτείνει τις δυνατότητες του πυρήνα του Joomla. Μέσω της ενότητα αυτής θα ενεργοποιήσουμε το reCapctha και θα εγκαταστήσουμε εναλλακτικά πρόσθετα.

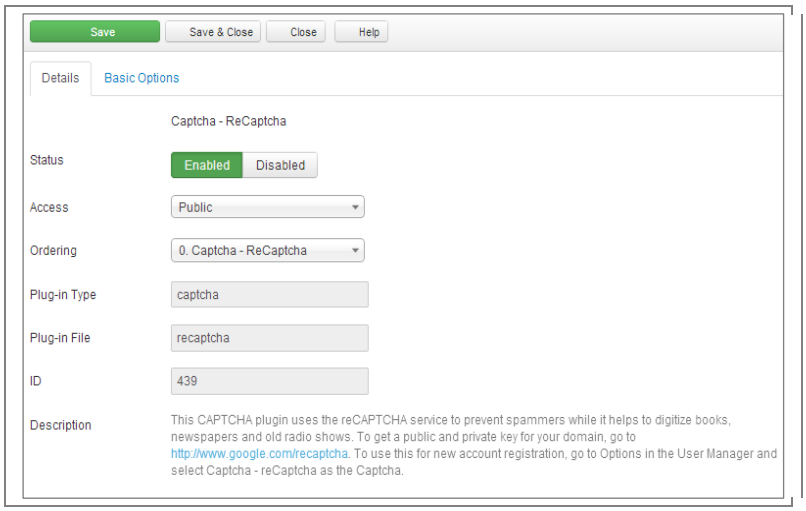
1.17.2.1. Ενεργοποίηση reCAPTCHA

Αφού έγιναν όλες οι απαραίτητες ενέργειες, που αφορούν το περιεχόμενο της σελίδας μας, στην ενότητα αυτή θα περιγράψουμε τον τρόπο ενεργοποίησης της προστασίας του περιεχομένου.

Η έκδοση του Jooma, που χρησιμοποιήσαμε έχει ενσωματωμένη την δυνατότητα ενεργοποίησης CAPTCHA και συγκεκριμένα το reCaptcha. Για τις ανάγκες της εργασίας στην επόμενη ενότητα θα εγκατασταθεί και το keyCaptcha, έτσι ώστε να έχουμε περισσότερες επιλογές, για να οδηγηθούμε σε ορθότερα συμπεράσματα, όσο αφορά την ασφάλεια που παρέχουν αυτά τα συστήματα.

Η ενεργοποίηση του ενσωματωμένου συστήματος reCaptcha γίνεται σε τέσσερα βήματα.

- **Βήμα 1^{ο} – Ενεργοποίηση του προσθέτου**: Μέσω της επιλογής “Extensions→Plugin Manager → Captcha-reCaptcha”.



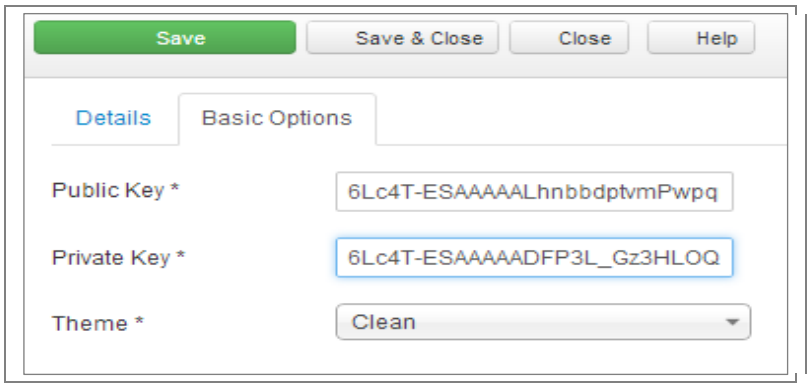
Εικόνα 3-30 Ενεργοποίηση reCaptcha

- **Βήμα 2^{ο} – Δημιουργία κλειδιών**: Για την δημιουργία των ιδιωτικών και δημόσιων κλειδιών που χρειάζονται θα πρέπει να επισκεφτούμε τον

αντίστοιχο ιστότοπο και να δηλώσουμε το όνομα τομέα (grindexes.gr). Τα κλειδιά δημιουργούνται αυτόματα και εμφανίζονται στην οθόνη μας.

Εικόνα 3-31 Δημιουργία κλειδιών reCaptcha

- **Βήμα 3^{ο} – Τοποθέτηση των κλειδιών**: Ενημέρωση του προσθέτου με τα κλειδιά που δημιουργήσαμε.



Εικόνα 3-32 Εισαγωγή κλειδιών reCaptcha

- **Βήμα 4^{ο} – Ενεργοποίηση reCaptcha για σχόλια**: Έλεγχος reCaptcha στα άρθρα του K2 για την αποφυγή σχολίων SPAM. Η ενεργοποίηση γίνεται μέσω "System→Global Configuration → Captcha-reCaptcha → K2", επιλέγοντας Enable reCaptcha in Comments.

Εικόνα 3-33 Ενεργοποίηση reCaptcha για σχόλια

1.17.2.2. Ενεργοποίηση Secureimage

Η εγκατάστασης του Captcha Securimage, γίνεται χειροκίνητα αφού δεν υπάρχει αντίστοιχο πρόσθετο για Joomla. Η χρήση του συγκεκριμένου Captcha, θα έχει επίδραση στην φόρμα εγγραφής νέων χρηστών.

Τα βήματα εγκατάστασης είναι:

- **Βήμα 1ο:** Μεταφόρτωση και αποσυμπίεση του λογισμικού στον ριζικό κατάλογο του ιστοτόπου μας “root \ securimage\”.
- **Βήμα 2ο:** Επεξεργασία του σεναρίου εγγραφής χρηστών το οποίο βρίσκεται στο κατάλογο “root \ templates\ [το πρότυπο που χρησιμοποιούμε] \ html \ com_K2 \ register.php”.
- **Βήμα 3ο**: Προσθήκη του κώδικα που ακολουθεί για την εμφάνιση του Secureimage Captcha. Για τη στηλοθέτηση των στοιχείων χρησιμοποιούμε ένα πίνακα, όπου εμφανίζουμε στη πρώτη στήλη το τίτλο, στη δεύτερη στήλη το πεδίο εισαγωγής της λέξης και στο τρίτο πεδίο την εικόνα Captcha.

```
<table class="admintable" cellpadding="0" cellspacing="0">
 <tr>
   <td class="key"> Captcha</td>
   <td>
     <input type="text" name="captcha_code" size="10" maxlength="6" />
     <a href="#" onclick="document.getElementById('captcha').src =
     '/securimage/securimage_show.php?' + Math.random(); return false"></a>
    <img id="captcha" src="securimage/securimage_show.php" alt="CAPTCHA Image" />
   </td>
 </tr>
</table>
```

1.17.2.3. Ενεργοποίηση Snaphost Captcha

Όπως είδαμε σε προηγούμενο βήμα χρησιμοποιήσαμε ένα άρθρο του Joomla, για να καλέσουμε ένα σενάριο που επιτρέπει την εκτέλεση μίας φόρμας επικοινωνίας. Για μεγαλύτερη ασφάλεια σε αυτή τη φόρμα θα ενσωματώσουμε το αντίστοιχο Captcha της SnapHost. Για να γίνει αυτό θα πρέπει πρώτα να δημιουργηθεί το

σενάριο της φόρμας με το όνομα contact_us.php, η οποία θα εμφανίζει την φόρμα ενώ παράλληλα θα περιλαμβάνει ένα κείμενο Captcha.

Στο πρώτο τμήμα του κώδικα, δηλώνονται οι παράμετροι του Captcha, όπως: Η διεύθυνση για τον έλεγχο και η διεύθυνση επιστροφής. Μπορούν επίσης να δηλωθούν και δευτερεύοντα στοιχεία όπως η αποστολή της IP διεύθυνσης του χρήστη ή αν θα σταλεί στο χρήστη μήνυμα επιβεβαίωσης.

```
<form action="http://www.SnapHost.com/captcha/send.aspx"
id="ContactUsCaptchaWebForm" method="post" onsubmit="return ValidateForm(this);">
<input name="skip_WhereToReturn" type="hidden" value= "../success.html" />
<input name="skip_Subject" type="hidden" value="Contact Us Form" />
```

Στο δεύτερο μέρος του κώδικα υπάρχουν δύο συναρτήσεις Javascript: η ValidateForm(frm) και η ReloadCaptchaImage(captchaImageId). Η πρώτη χρησιμοποιείται για έλεγχο των πεδίων κατά την αποστολή της φόρμας, ενώ η δεύτερη φορτώνει μία νέα εικόνα Captcha στη περίπτωση που πατηθεί το κουμπί ανανέωσης.

```
function ValidateForm(frm) {
    if (frm.Name.value == "")
   {
     alert('Το όνομα απαιτείται');
     frm.Name.focus();return false;
    }
   if (frm.FromEmailAddress.value == "")
   {
     alert('Η διεύθυνση email είναι υποχρεωτική.');
     frm.FromEmailAddress.focus();
     return false;
   }
  if (frm.FromEmailAddress.value.indexOf("@") <1|| frm.FromEmailAddress.value.indexOf(".") < 1)
   {
     alert('Δώστε μια έγκυρη διεύθυνση');
     frm.FromEmailAddress.focus();return false;
   }
    if (frm.Comments.value == "")
   {
     alert('Η αποστολή κειμένου είναι υποχρεωτική.');
     frm.Comments.focus();
     return false;
    }
    if (frm.skip_CaptchaCode.value == "")
   {
     alert('Πληκτρολογήστε τον κωδικό που βλέπετε');
     frm.skip_CaptchaCode.focus();
     return false;
   }
```

```
  return true;
}
function ReloadCaptchaImage(captchaImageId) {
    var obj = document.getElementById(captchaImageId);
    var src = obj.src;
    var date = new Date();
    var pos = src.indexOf('&rad=');
    if (pos >= 0)
    {
       src = src.substr(0, pos);
    }
    obj.src = src + '&rad=' + date.getTime();
    return false;
}
```

Το τελευταίο μέρος του κώδικα, σχεδιάζεται η φόρμα με τα αντίστοιχα πεδία και γίνεται ο έλεγχος και καλούνται τα αντίστοιχα σενάρια javascript.

```
<table border="0" cellpadding="5" cellspacing="0" width="600">
    <tr>
        <td><b>Όνομα*:</b></td>
        <td><input name="Όνομα" type="text" maxlength="60" style="width:350px;" /></td>
    </tr>
    <tr>
        <td><b>Επώνυμο:</b></td>
        <td><input name="Τηλέφωνο" type="text" maxlength="43" style="width:350px;" /></td>
    </tr>
    <tr>
        <td><b>Email*:</b></td>
        <td><input name="Email" type="text" maxlength="60" style="width:350px;" /></td>
    </tr>
    <tr>
        <td><b>Σχόλια*:</b></td>
        <td><textarea name="Σχόλια" rows="7" cols="40" style="width:350px;">
        </textarea></td>
    </tr>
    <tr>
        <td colspan="2" align="center"> <br />
            <table border="0" cellpadding="0" cellspacing="0">
                <tr valign="top">
                    <td> <i>Πληκτρολογήστε τον κωδικό*:</i>
                        <input name="skip_CaptchaCode"
                                type="text" style="width:80px;" maxlength="6" />
                    </td>
                    <td>
                        src="http://www.SnapHost.com/captcha/CaptchaImage.aspx?
                        id=GPKK4JNMTMH8&ImgType=2" /></a><br />
                        <a href="#" onclick="return ReloadCaptchaImage('CaptchaImage');">
                        <span style="font-size:12px;">Επαναφόρτωση</span></a>
                    </td>
                </tr>
            </table> <br />
            * - απαιτούμενα πεδία.              
            <input name="skip_Submit" type="submit" value="Submit" />
        </td>
```

```
    </tr>
</table><br />
</form>
```

Το αρχείο success.html, αντίστοιχα ενημερώνει το χρήστη αν πέρασε με επιτυχία το τεστ captcha και ανακατευθύνει τον χρήστη μετά την αποστολή του μηνύματος στην αρχική οθόνη μετά από 5 δευτερόλεπτα.

```
<html>
  <head>
    <script type="text/javascript">
      function delayer(){window.location = "http://127.0.0.1/captcha/"          }
    </script>
  </head>
  <body onLoad="setTimeout('delayer()', 5000)">
     <p>Ευχαριστούμε για το μήνυμά σας!</p> <h2>Ανακατεύθυνση</h2> <p>Μετά από 5 δευτερόλεπτα θα γίνει ανακατεύθυνση στην αρχική σελίδα!</p>
   </body>
</html>
```

1.17.2.4. Ενεργοποίηση Opencaptcha

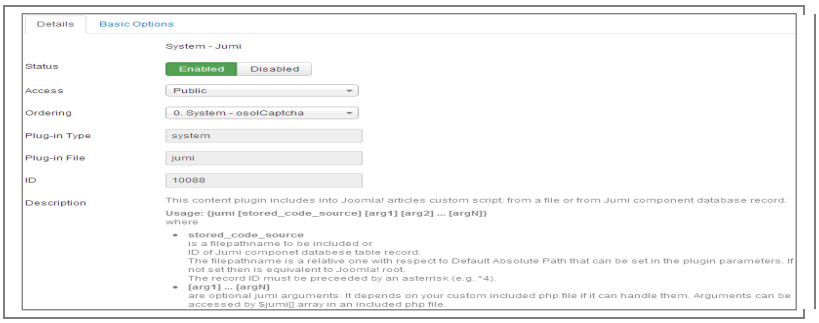
Για τις ανάγκες τις εργασίας έχουμε εγκαταστήσει μία σειρά Captcha συστημάτων από διαφορετικές πηγές. Το τελευταίο Captcha που θα εξετάσουμε είναι το Opencaptcha και θα χρειαστεί τροποποίηση του κώδικα του Joomla, για να το ενεργοποιήσουμε. Συγκεκριμένα θα το ενσωματώσουμε στην φόρμα εισόδου, τροποποιώντας το αντίστοιχο αρχείου που είναι: “[ριζικός κατάλογος] \ templates \ [πρότυπο που χρησιμοποιούμε] \ html \ com_users \ login”

1.17.2.5. Εγκατάσταση Jumi

Η συγγραφή κώδικα, μέσω του περιεχομένου δεν είναι μία δυνατότητα ενσωματωμένη στον πυρήνα του Joomla. Η κλήση από μία φόρμα, ή μέσω ενός άρθρου κώδικα php, javascript ή ακόμα και κάποιου εξωτερικού σεναρίου, απαιτεί την επέμβαση (hacking), του κώδικα του Joomla. Άλλη μια λύση είναι η χρήση προσθέτων που επιτρέπουν αυτή την δυνατότητα, διευκολύνοντας έτσι τον διαχειριστή για τις απαιτούμενες ενέργειες. Στην παρούσα εργασία, θα δείξουμε και τους δύο τρόπους.

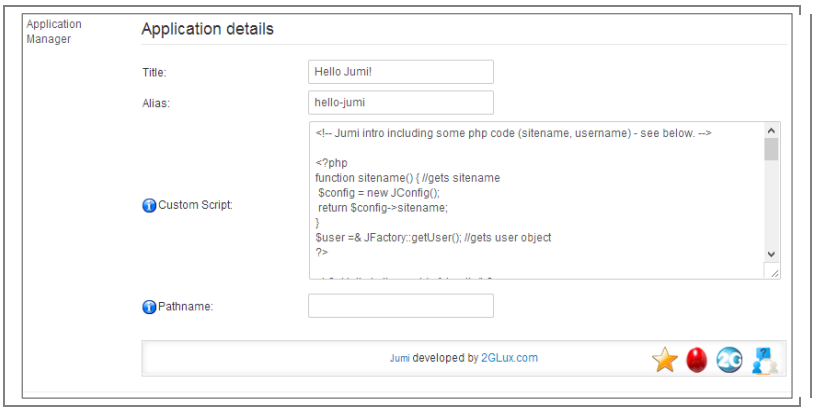
Όσο αφορά τον τρόπο που απαιτεί την χρήση κάποιου εξωτερικού προσθέτου, θα εγκαταστήσουμε το Jumi το οποίο επιτελεί αυτό τον σκοπό. Το συστατικό jumi, είναι ένα πρόσθετο το οποίο το εγκαθιστούμε μέσα από τη διαχείριση προσθέτων. Ουσιαστικά επεκτείνει την λειτουργία των άρθρων του Joomla, επιτρέποντας την εκτέλεση κώδικα μέσα από αυτά. Μετά την εγκατάσταση, το Jumi μπορεί να χρησιμοποιηθεί με δύο τρόπους:

- **Ως πρόσθετο:** Ενεργοποιούμαι ως πρόσθετο και κάθε φορά που αναφερόμαστε με την αντίστοιχη οδηγία, αυτόματα εκτελεί τον κώδικα.



Εικόνα 3-34 Πρόσθετο jumi

- **Ως εφαρμογή:** Δημιουργούμε αντίστοιχα σενάρια κώδικα και τα ενεργοποιούμε μέσα από μία επιλογή μενού.



Εικόνα 3-35 Εφαρμογή jumi

1.17.3. Περιεχόμενο

Η δυνατότητα εμφάνισης περιεχομένου μέσω Joomla, δίνεται κυρίως μέσω της ενότητας των άρθρων. Αυτή όμως η επιλογή έχει κάποιους περιορισμούς όπως, η αδυναμία δημιουργίας συγκεκριμένων πεδίων. Για την δημιουργία ενός ιστοτόπου ευρετηρίασης, χρειαζόμαστε να ορίσουμε τα δικά μας πεδία όπως για παράδειγμα: Πόλη, Διεύθυνση καταστήματος, Τηλέφωνο επικοινωνίας, Ηλεκτρονική διεύθυνση επικοινωνίας, Url κλπ.

1.17.3.1. CCK K2

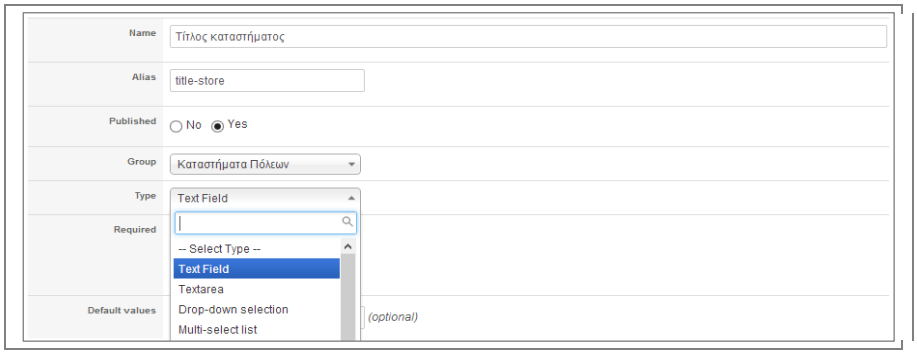
Ο τρόπος για να το πετύχουμε είναι με την εγκατάσταση ενός προσθέτου κατασκευής περιεχομένου (CCK: Content Construction Kit). Από τα πιο διαδεδομένα

είναι το K2 με πολλές δυνατότητες. Έχει πολύ καλή τεκμηρίωση και είναι λογισμικό ανοιχτού κώδικα. Η εγκατάσταση γίνεται πάλι μέσω του Διαχειριστή Επεκτάσεων.

Για την δημιουργία ενός άρθρου σε K2, το πρώτο πράγμα που πρέπει να γίνει, είναι η προσθήκη των νέων πεδίων που θέλουμε. Πρώτα όμως, πρέπει να ορίσουμε την κατηγορία αυτών των πεδίων, επιλέγοντας από την οθόνη διαχείρισης του K2 «Extra Field Groups». Στη συνέχεια επιλέγουμε «Νέο» και πληκτρολογούμε το όνομα της κατηγορίας και αποθηκεύουμε. Εφόσον θέλουμε να ορίσουμε πεδία που αφορούν την παρουσίαση των καταστημάτων θα ονομάσουμε την κατηγορία «Καταστήματα Πόλεων»

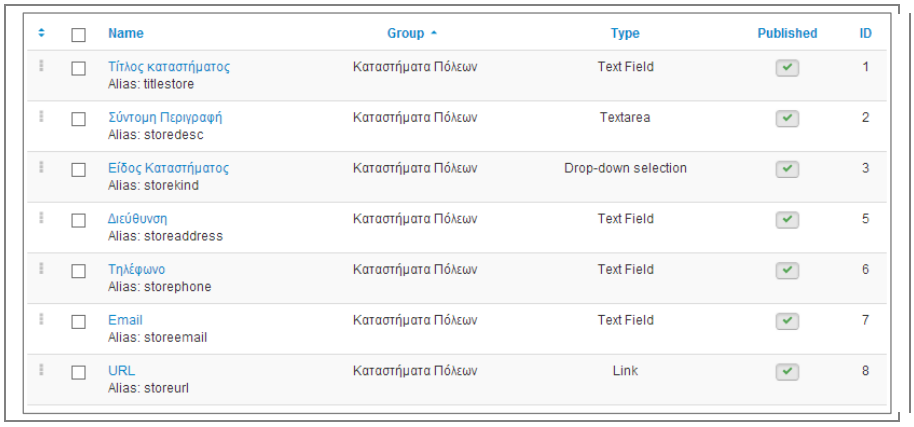
Εικόνα 3-36 Δημιουργία κατηγορίας πρόσθετων πεδίων

Στη συνέχεια θα δημιουργήσουμε τα αντίστοιχα πεδία μέσω της επιλογής “Extra Fields”. Εκεί θα ορίσουμε το όνομα του πεδίου, σε ποια κατηγορία ανήκει καθώς και τον τύπο του.



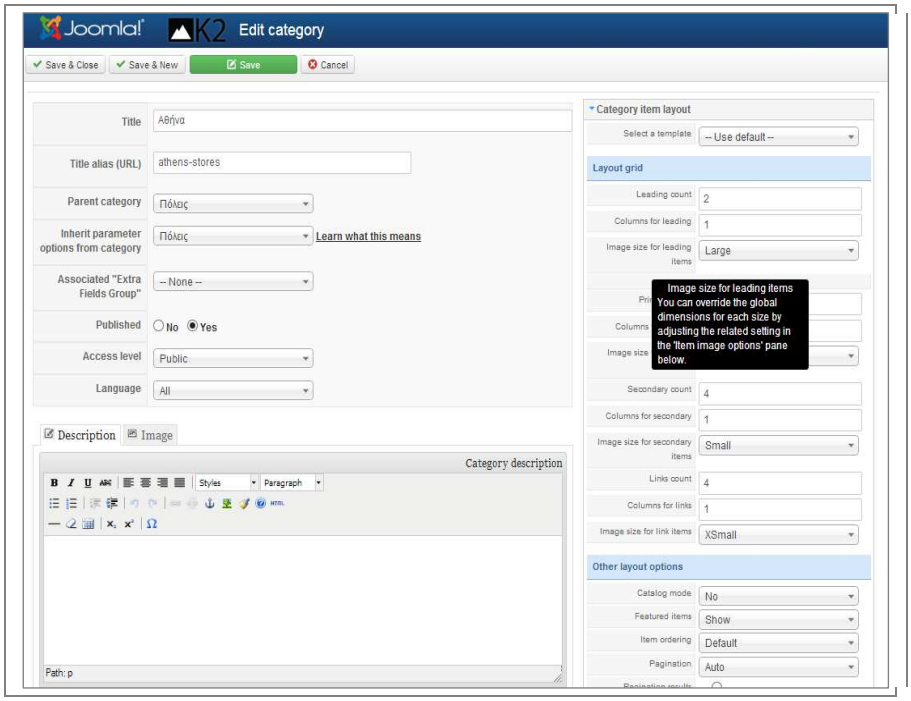
Εικόνα 3-37 Δημιουργία πρόσθετων πεδίων

Τα πεδία που θα δημιουργήσουμε καθώς και οι τύποι τους είναι:



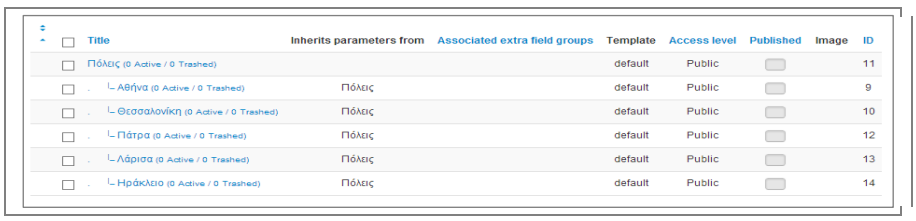
Εικόνα 3-38 Πρόσθετα πεδία

Την ίδια διαδικασία επιλέγουμε για όλες τις κατηγορίες άρθρων. Στη δική μας περίπτωση, θα χρησιμοποιήσουμε μία ριζική κατηγορία «Πόλεις» και κάθε πόλη θα αποτελεί υπό-κατηγορία.



Εικόνα 3-39 Δημιουργία κατηγορίας

Μετά την ολοκλήρωση της διαδικασίας έχουμε τις ακόλουθες κατηγορίες:

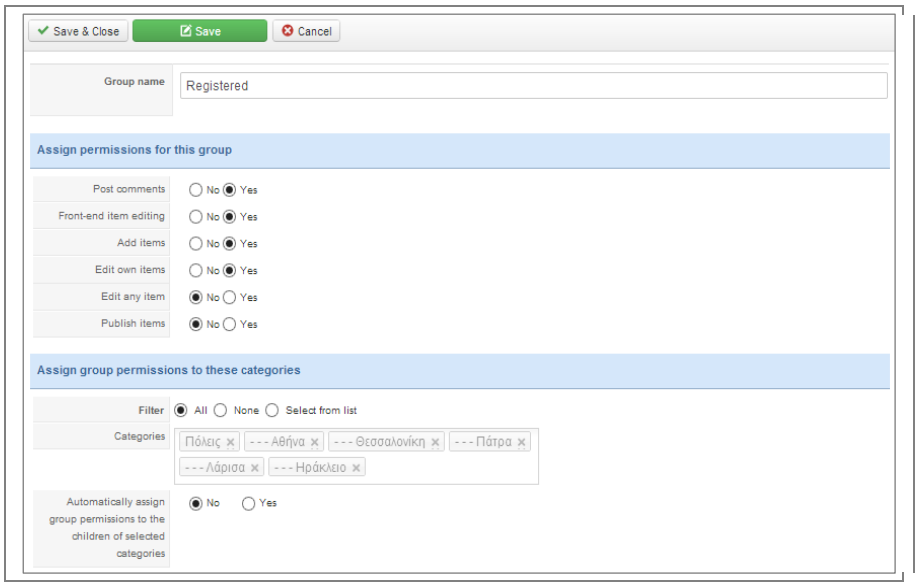


Εικόνα 3-40 Κατηγορίες K2

Αντίστοιχα θα μπορούσαμε να έχουμε ορίσει μία δεύτερη ριζική κατηγορία με τα είδη καταστημάτων. Επίσης έχουμε τη δυνατότητα τρίτου επιπέδου υπό-κατηγοριών, που θα μπορούσε για παράδειγμα να ήταν το είδος των καταστημάτων ανά πόλη.

Η υποστήριξη κύριων κατηγοριών και υπό-κατηγοριών μας εξασφαλίζει την εύκολη επέκταση των δυνατοτήτων του ιστοτόπου μας, όπως για παράδειγμα την δημιουργία ριζικής κατηγορίας ανά είδος καταστήματος απ' ευθείας.

Για την καταχώριση των καταστημάτων θα πρέπει να ορίσουμε δικαιώματα. Για να μπορεί κάποιος να κάνει καταχώριση θα πρέπει να είναι εγγεγραμμένος χρήστης στο σύστημα. Εφόσον αυτό ισχύει, τότε θα μπορεί να καταχωρεί και να επεξεργάζεται ένα κατάστημα. Ο ορισμός των δικαιωμάτων γίνεται μέσω της επιλογής "Components → K2 → User Groups". Η ομάδα "Registered" αφορά τους εγγεγραμμένους χρήστες, συνεπώς ο ορισμός των δικαιωμάτων θα επηρεάσει και τους χρήστες που συμμετέχουν.



Εικόνα 3-41 Δικαιώματα χρηστών

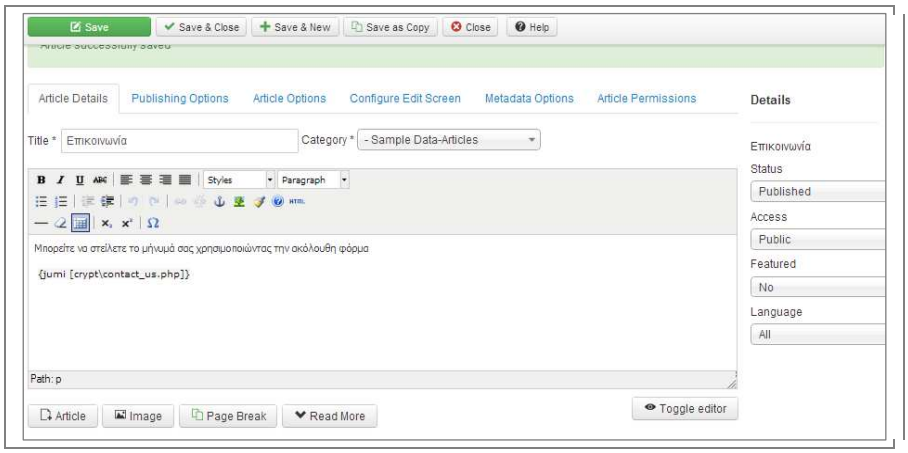
1.17.3.2. Άρθρα Joomla

Μέσα από αυτή την ενότητα, θα δημιουργήσουμε ένα άρθρο στο οποίο θα ενσωματώσουμε μία φόρμα επικοινωνίας. Για μεγαλύτερη ασφάλεια στη συνέχεια θα ενεργοποιήσουμε και ένα μηχανισμό Captcha για τη συγκεκριμένη φόρμα.

Η δημιουργία των άρθρων γίνεται μέσω της επιλογής "Content → Article Manager → New Article". Μπορούμε να ομαδοποιήσουμε τα άρθρα σε κατηγορίες, αλλά στην

περίπτωσή μας δεν χρειάζεται εφόσον πρόκειται για τη δημιουργία μιας φόρμας επικοινωνίας.

Εκτός από τον τίτλο του άρθρου, σημαντικό πεδίο αποτελεί το περιεχόμενο, το οποίο θα εμφανίζεται στην σχετική σελίδα του ιστοτόπου μας.



Εικόνα 3-42 Δημιουργία άρθρου

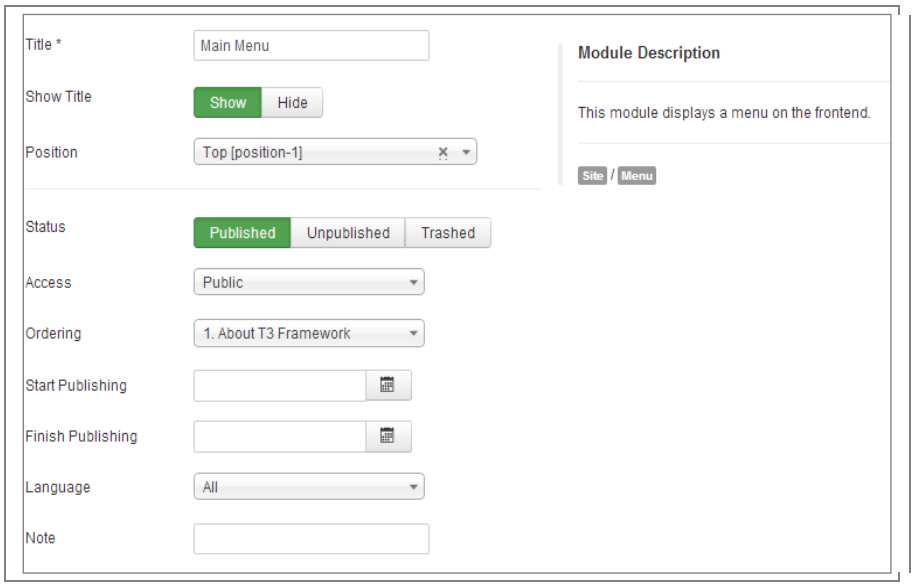
Στη περίπτωσή μας επιλέγουμε να εμφανίσουμε ένα μήνυμα στον επισκέπτη του ιστοτόπου, ενώ παράλληλα καλούμε το σενάριο που έχουμε δημιουργήσει, μέσω του κώδικα {jumi[crypt\contact_us.php]}. Η οδηγία jumi χρησιμοποιείται για να δηλώσουμε, ότι αυτό που θα ακολουθήσει είναι μία κλήση σε μία συνάρτηση ή αρχείο.

1.17.4. Μενού Επιλογών

Το μενού επιλογών αποτελεί το σύστημα πλοήγησης στον ιστότοπό μας, καθώς και τον τρόπο αλληλεπίδρασης του χρήστη με αυτόν. Εκτός από την πλοήγηση, θα μπορεί ο επισκέπτης να εκτελεί και εργασίες όπως:

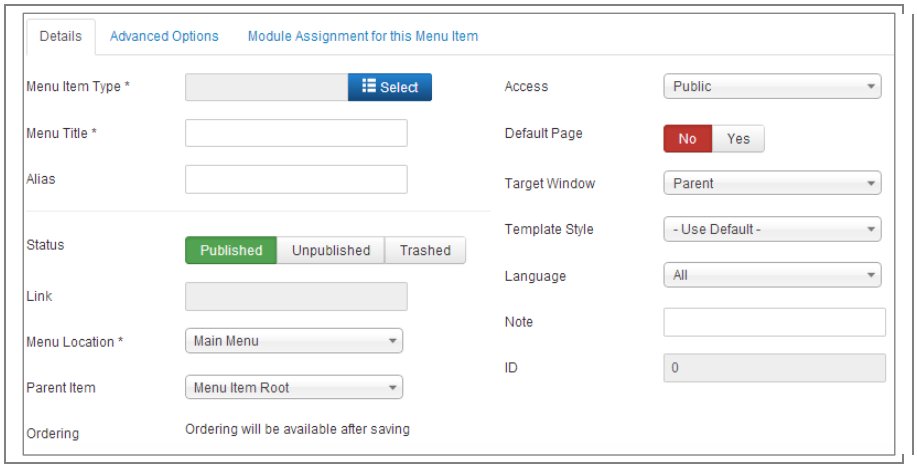
- Εγγραφή στο σύστημα
- Σύνδεση στο σύστημα
- Καταχώριση καταστήματος
- Εμφάνιση κατηγοριών
- Φόρμα επικοινωνίας
- Επικοινωνία

Για την δημιουργία του μενού και των στοιχείων του, απαραίτητη προϋπόθεση είναι να έχει δημιουργηθεί το αντίστοιχο ένθεμα (module). Αυτό γίνεται μέσα από την επιλογή “Extensions → Module Manager → New → Menu”. Σε αυτή την σελίδα καθορίζουμε τα βασικά του χαρακτηριστικά όπως όνομα, θέση εμφάνισης, αν θα είναι ενεργό κλπ.



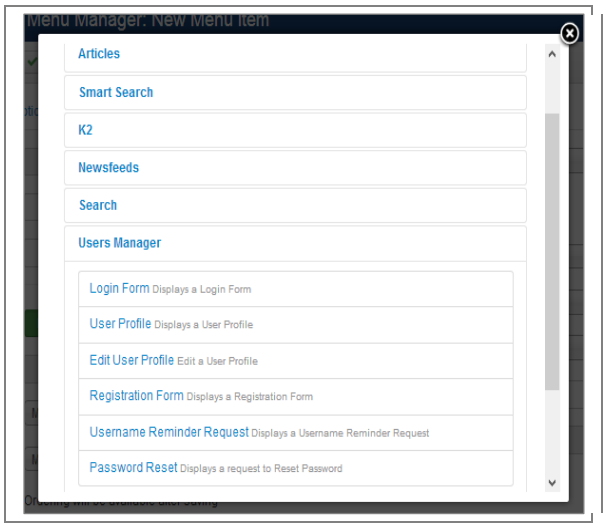
Εικόνα 3-43 Δημιουργία Μενού

Μετά την δημιουργία του ενθέματος, θα πρέπει να οριστούν τα στοιχεία του μενού. Αυτό γίνεται μέσω της επιλογής “Menus → [Όνομα μενού] → “New”, που στη περίπτωσή μας είναι το όνομα μενού είναι “Main Menu”.



Εικόνα 3-44 Δημιουργία στοιχείου μενού

Η πρώτη επιλογή για τη δημιουργία στοιχείου μενού, είναι η πιο σημαντική γιατί καθορίζει το τύπο του στοιχείου. Πατώντας “Select”, οδηγούμαστε σε μια δευτερεύουσα φόρα με τους διαθέσιμους τύπους.

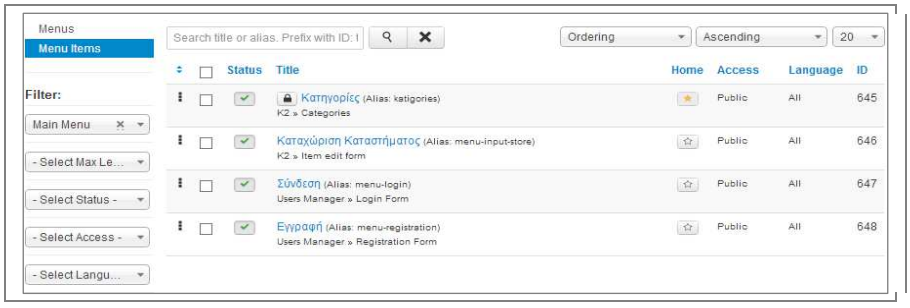


Εικόνα 3-45 Τύποι στοιχείων

Συνεπώς για τις επιλογές του δικού μας μενού οι αντίστοιχοι τύποι είναι:

- **Εγγραφή στο σύστημα:**Users Manager: Registration Form για την εμφάνιση της φόρμας εγγραφής.
- **Σύνδεση στο σύστημα**: Users Manager: Login Form για την εμφάνιση της φόρμας εισόδου.
- **Καταχώριση καταστήματος**: K2-Item:Edit Form, για την εμφάνιση φόρμας καταχώρισης καταστήματος.
- **Εμφάνιση κατηγοριών**: K2:Caterories για την εμφάνιση των κατηγοριών των καταστημάτων.

Η εικόνα που ακολουθεί δείχνει την δομή του μενού όπως διαμορφώθηκε μετά την δημιουργία των στοιχείων.



Εικόνα 3-46 Δομή Μενού

Το στοιχείο “Κατηγορίες” αντιστοιχεί στο “Home”, που σημαίνει ότι είναι η εξ’ ορισμού επιλογή όταν κάποιος χρήστης επισκεφθεί τον ιστότοπο.

1.18. ΠΑΡΟΥΣΙΑΣΗ ΙΣΤΟΤΟΠΟΥ (backend)

Οι εργασίες που έγιναν στην προηγούμενη ενότητα στον back-end του Joomla, έχουν ορατά αποτελέσματα στο front-end. Αυτός είναι και ο κύριος μηχανισμός λειτουργίας των συστημάτων διαχείρισης περιεχομένου.

1.18.1. Αρχική Σελίδα - Κατηγορίες

Η αρχική μας σελίδα αποτελείται, φαίνεται στην εικόνα που ακολουθεί.

Εικόνα 3-47 Αρχική σελίδα

Αποτελείται από τα στοιχεία:

- **K2 Category Listing (1)**: Εδώ εμφανίζονται οι υπο-κατηγορίες των στοιχείων του K2, δηλαδή η κάθε πόλη. Επιλέγοντας μία υπό-κατηγορία εμφανίζονται τα καταστήματα της συγκεκριμένης πόλης. Οδηγούμαστε σε αυτή την οθόνη επιλέγοντας "Κατηγορίες" από το μενού (2).
- **Μενού επιλογών (2):** Είναι το μενού επιλογών (Main Menu) που δημιουργήσαμε στο προσκήνιο. Η κάθε επιλογή μας οδηγεί σε διαφορετική σελίδα.
- **Λογότυπο (3):** Το λογότυπο της σελίδας μας.
- **Αναζήτηση (4):** Γίνεται αναζήτηση στο περιεχόμενο του ιστοτόπου.
- **Τελευταίες καταχωρίσεις (5):** Εμφανίζει τις τελευταίες καταχωρίσεις καταστημάτων.

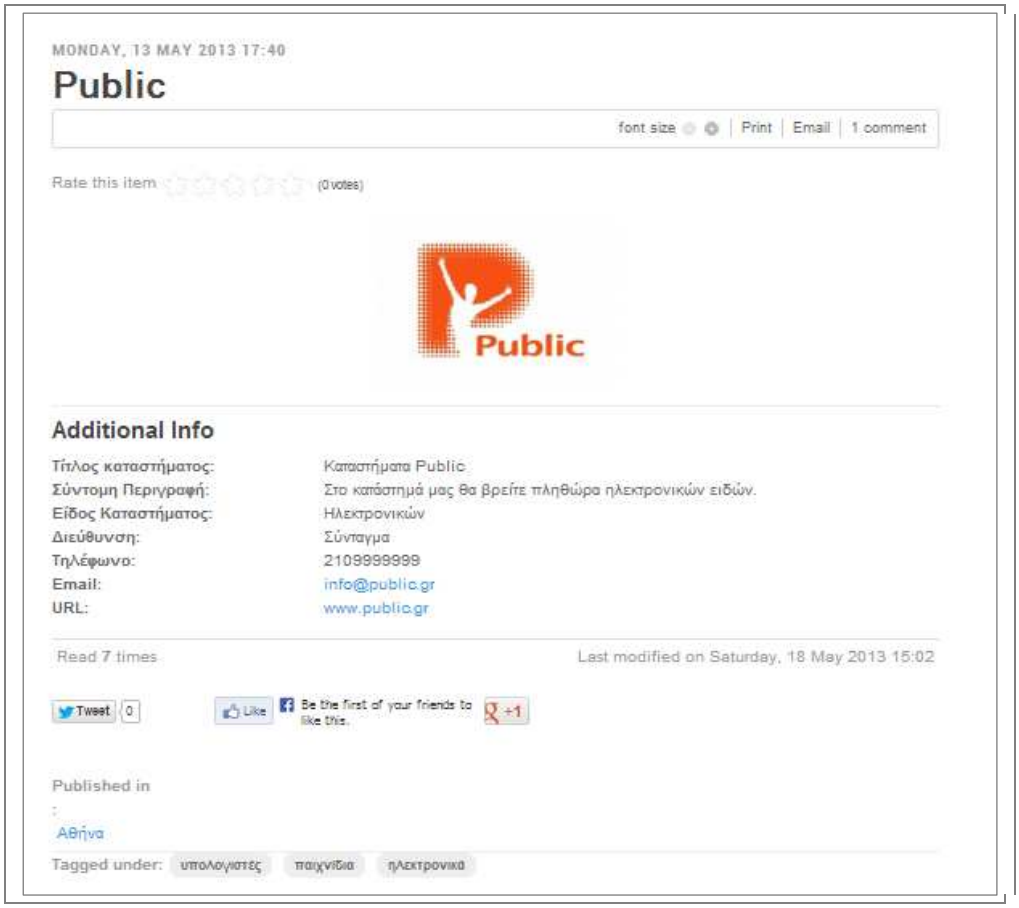
1.18.2. Σελίδα πόλεων

Στην σελίδα των πόλεων, μπορούμε να βρεθούμε επιλέγοντας την αντίστοιχη πόλη, από την αρχική σελίδα. Εδώ βλέπουμε μία λίστα καταστημάτων, για την συγκεκριμένη πόλη. Σε κάθε κατάστημα υπάρχει το λογότυπο, η κατηγορία που ανήκει και οι σχετικές θεματικές ετικέτες (tags). Πατώντας σε ένα κατάστημα οδηγούμαστε στην σελίδα του. Αντίστοιχα πατώντας σε μία ετικέτα επιστρέφονται όλα τα καταστήματα που περιέχουν αυτή την ετικέτα.

Εικόνα 3-48 Σελίδα πόλεων

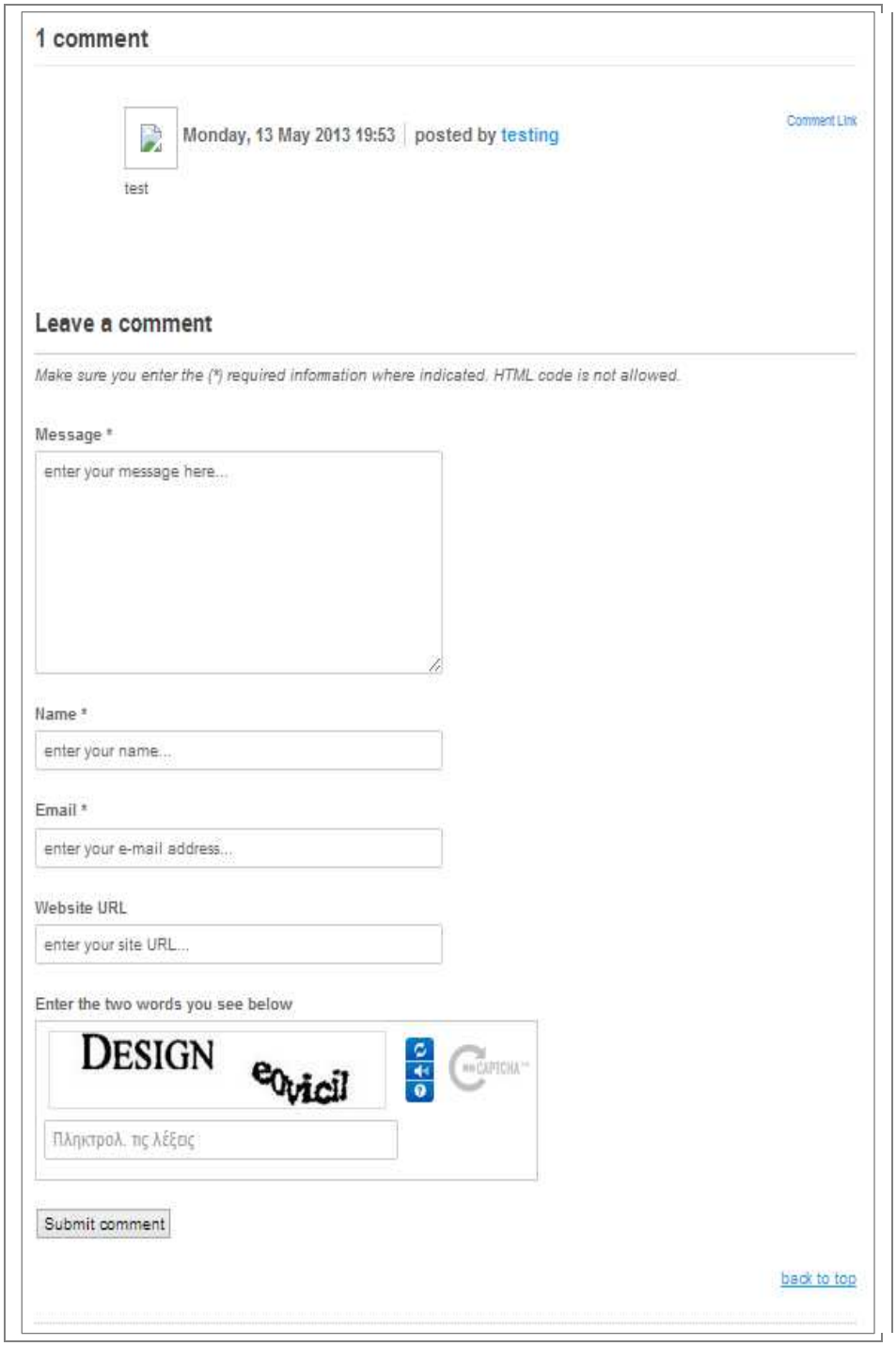
1.18.3. Σελίδα καταστήματος

Η σελίδα καταστήματος, εμφανίζει αναλυτικές πληροφορίες σχετικά με το κατάστημα όπως το λογότυπο, επιπρόσθετες πληροφορίες που αντιστοιχούν στα εξτρά πεδία που δημιουργήσαμε μέσω του K2 και ετικέτες. Αντίστοιχα υπάρχουν και τα κουμπιά κοινοποίησης του περιεχομένου στα κοινωνικά δίκτυα.



Εικόνα 3-49 Σελίδα καταστήματος

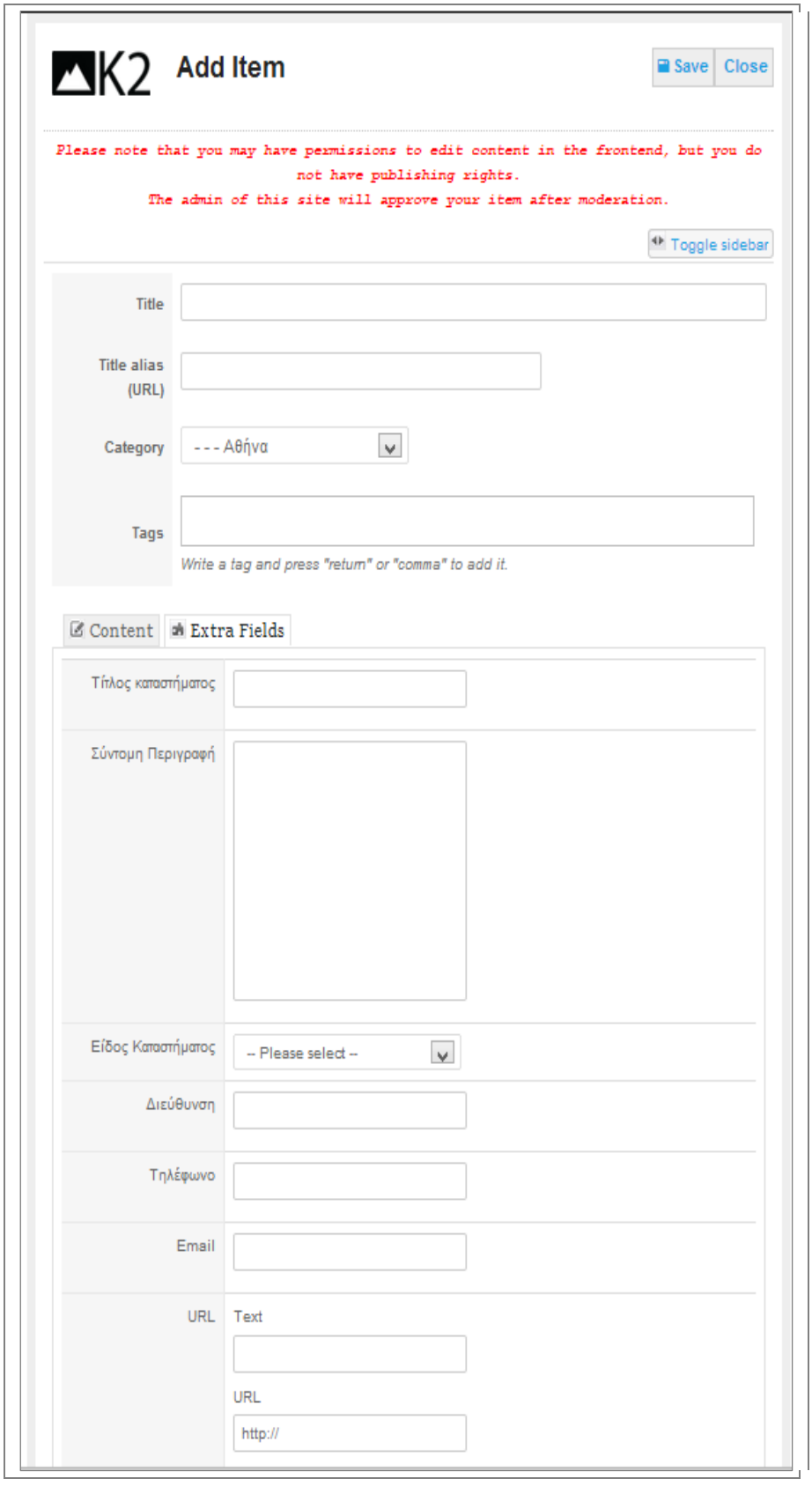
Μετά το τμήμα της σελίδας που αναφέρεται στα στοιχεία του καταστήματος, υπάρχει η δυνατότητα του επισκέπτη να σχολιάσει. Η δυνατότητα αυτή για λόγους ασφάλειας και αποτροπής ανεπιθύμητων σχολίων, έχει περιοριστεί μόνο με τη χρήση captcha. Το captcha, που χρησιμοποιείται είναι reCaptcha, όπως φαίνεται στην εικόνα.



Εικόνα 3-50 Εισαγωγή σχολίων με τον έλεγχο reCaptcha

1.18.4. Καταχώριση καταστήματος

Η προσπέλαση της φόρμας για την καταχώριση καταστήματος γίνεται μέσα από το μενού της αρχικής σελίδας. Η καταχώριση, γίνεται μέσω της φόρμας της εικόνας 3-28.

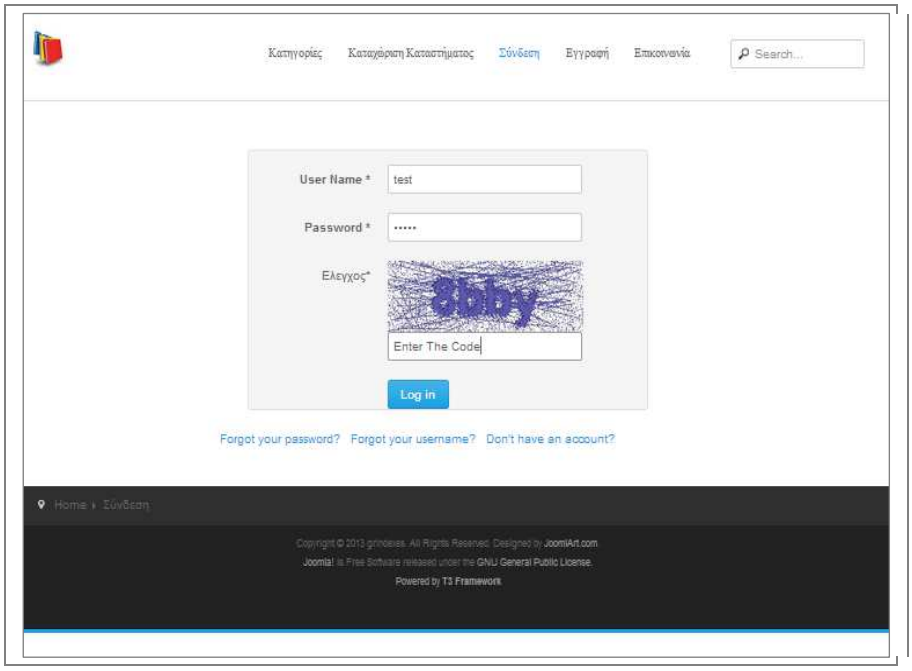


Εικόνα 3-51 Καταχώριση καταστήματος

Η καταχώριση επιτρέπεται μόνο από εγγεραμμένους χρήστες στο σύστημα, συνεπώς δεν θεωρήθηκε σκόπιμη η χρήση περαιτέρω μέτρων προστασίας.

1.18.5. Σύνδεση χρήστη

Η σύνδεση, επιτρέπει σε ένα χρήστη να ταυτοποιηθεί για να αποκτήσει δικαιώματα καταχωρητή καταστήματος. Πολλές φορές οι οθόνες σύνδεσης, χρησιμοποιούνται από κακόβουλους χρήστης για τις λεγόμενες brute-force attack επιθέσεις. Σε αυτές τις επιθέσεις ένας χρήστης με ένα πρόγραμμα, προσπαθεί από ένα λεξικό συνθηματικών να αποκτήσει πρόσβαση στην σελίδα. Ένας επιπλέον επίπεδο ασφάλειας είναι η χρήση captcha. Στην σελίδα της σύνδεσης έχει χρησιμοποιηθεί το opencaptcha.

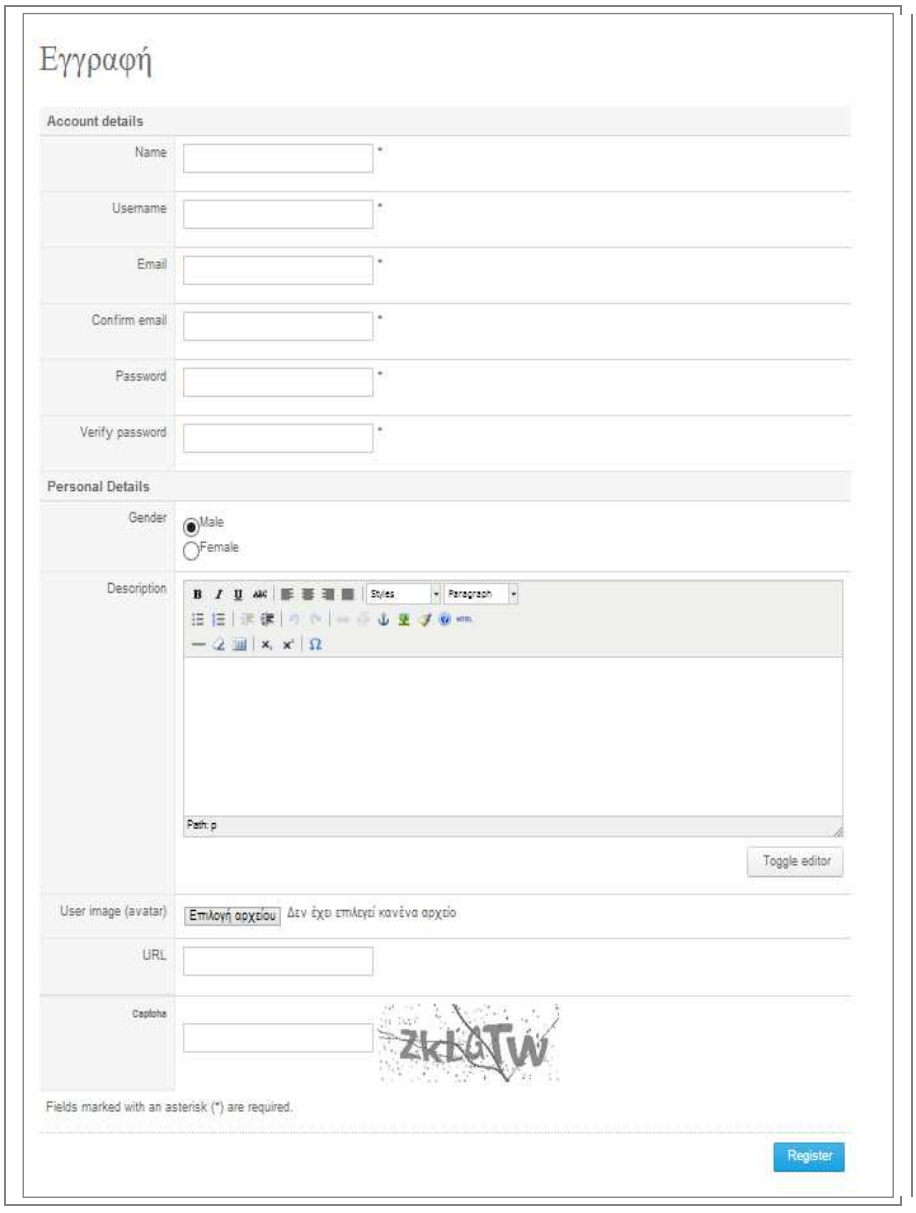


Εικόνα 3-52 Χρήση opencaptcha στην σελίδα της σύνδεσης

1.18.6. Εγγραφή χρήστη

Αντίστοιχα με την σύνδεση, κακόβουλοι χρήστες μπορεί να χρησιμοποιούν προγράμματα για την αυτόματη εγγραφή χρηστών. Αυτό μπορεί για παράδειγμα να χρησιμοποιηθεί από υπηρεσίες black hat seo, για την δημιουργία προφίλ χρηστών παρέχοντας backlinks. Ένα backlink είναι ένας σύνδεσμος, ο οποίος δείχνει έναν ιστότοπο για μία λέξη κλειδί. Αν για τη συγκεκριμένη λέξη κλειδί, υπάρχουν πολλοί

σύνδεσμοι σε αυτό τον ιστότοπο, τότε αυτός αποκτά σπουδαιότητα στα αποτελέσματα των μηχανών αναζήτησης. Συνεπώς η δημιουργία προφίλ με συνδέσμους, εξυπηρετεί αυτή την τεχνική.



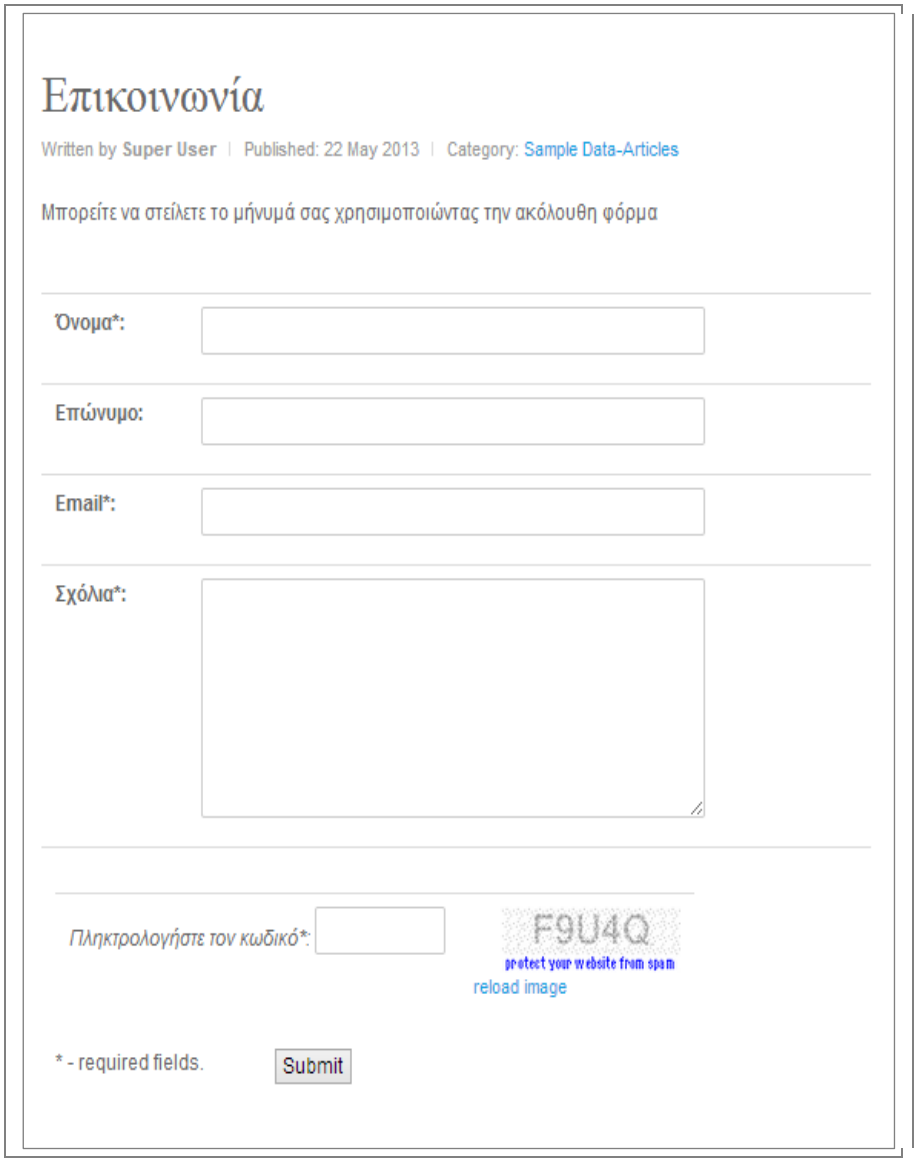
Εικόνα 3-53 Φόρμα εγγραφής

Στη φόρμα της εγγραφής, ο χρήστης δίνει τα απαραίτητα στοιχεία, ενώ ο έλεγχος captcha γίνεται με τη χρήση του secureimage.

1.18.7. Επικοινωνία

Η φόρμα της επικοινωνίας, χρησιμοποιείται από τους επισκέπτες για την αποστολή μηνυμάτων προς τον υπεύθυνο της σελίδας. Μπορεί να χρησιμοποιηθεί, όμως για την

αποστολή μαζικών ανεπιθύμητων μηνυμάτων. Έτσι σε αυτή την σελίδα, έχουμε χρησιμοποιήσει το captcha της snaphost.



Εικόνα 3-54 Φόρμα επικοινωνίας

ΠΗΓΕΣ ΚΕΦΑΛΑΙΟΥ

- *Content Management Systems - The History and the Future.* (2013, 05 20). Ανάκτηση από Ezine Articles: http://ezinearticles.com/?Content-Management-Systems---The-History-and-the-Future&id=1665607.

- *Content Management System*: Ανάκτηση (2013, 05 20) από Wikipedia: http://en.wikipedia.org/wiki/Content_management_system.

- *Joomla 1.5 API Reference*. (2011,06 13). Ανάκτηση από Joomla: http://api.joomla.org/li_Joomla-Framework.html

- *Model-view-controller*. (2011,05 20). Ανάκτηση από Wikipedia: http://en.wikipedia.org/wiki/Model-view-controller.

- Porst, T. (2009). *Joomla! 1.5 content administration* . S.l.: Packet Pub Ltd.

- Severdia, R., & Crowder, K. (2010). *Using Joomla* . Sebastopol, Calif.: O'Reilly Media.

- *GNU General Public Licence*. (2013, 06 03). Ανάκτηση από Wikipedia: http://el.wikipedia.org/wiki/GNU_General_Public_License

- *XAMPP*. (2013, 05 03). Ανάκτηση από ApacheFriends: http://www.apachefriends.org/en/xampp.html

- Mariott, J., & Waring, E. (2010). pp.174, 183-184. *The Official Joomla! Book.* Addison Wesley.

- JoomlaWorks (2013, 05 03). Ανάκτηση από Simple Image Gallery: http://www.joomlaworks.gr/content/view/17/42/

– ΠΑΡΑΚΑΜΨΗ CAPTCHA

1.19. ΓΕΝΙΚΑ

Από την περίοδο εμφάνισης των Captcha, άρχισαν αντίστοιχα οι προσπάθειες μηχανικής επίλυσής τους. Τα συστήματα επίλυσης, χρησιμοποιούν αλγόριθμους, σχεδιασμένους να εξάγουν μεμονωμένα σύμβολα, χρησιμοποιώντας την τεχνολογία της οπτικής αναγνώρισης (OCR).

Παρ' όλα αυτά, η δημιουργία αυτών των αλγορίθμων είναι μια πολύπλοκη διαδικασία, ενώ το αποτέλεσμα πολλές φορές δεν είναι ικανοποιητικό. Στη συνέχεια θα παρουσιάσουμε από αυτές τις προσπάθειες.

1.20. ΤΕΧΝΟΛΟΓΙΕΣ

Οι υπηρεσίες επίλυσης κυρίως, χρησιμοποιούν ανεπτυγμένη τεχνολογία OCR, για την αποκωδικοποίηση των Captcha. Σε μερικές περιπτώσεις τα Captcha είναι δύσκολο να επιλυθούν, ιδιαίτερα όταν δεν είναι κειμενικά. Για αυτό το λόγο, οι εταιρείες που προσφέρουν αυτές τις υπηρεσίες, προχωρούν ένα βήμα παραπέρα και προσλαμβάνουν ανθρώπινο δυναμικό, με χαμηλό κόστος (Motoyama et.al, 2012). Μία άλλη τεχνική είναι η παράκαμψη του Captcha, όχι με επίλυση αλλά με τεχνικές παράκαμψης.

1.20.1. OCR

Η οπτική αναγνώριση χαρακτήρων (OCR: Optical Character Recognition), ορίζεται ως η μηχανική ή ηλεκτρονική μετατροπή, σαρωμένων κειμένων σε αναγνώσιμη μορφή από έναν υπολογιστή. Με αυτό τον τρόπο αποφεύγεται η επαναπληκτρολόγηση μεγάλων κειμένων.

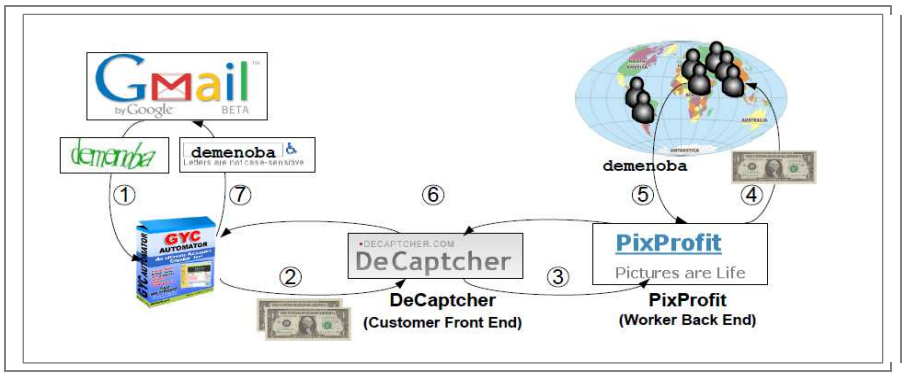
Υπάρχουν δυο βασικοί τρόποι ανίχνευσης τέτοιων κειμένων: Η πρώτη μέθοδος χρησιμοποιεί αντιστοίχιση με πρότυπα, ενώ η δεύτερη γίνεται με εξαγωγή των χαρακτήρων.

1.20.2. HUMAN SOLVING

Η επίλυση γρίφων CAPTCHA αποτελεί στις μέρες ένα πολύ έντονο οικονομικό πεδίο δράσης. Το οικοσύστημα που έχει δημιουργηθεί δημιουργεί κέρδη εκατομμυρίων στις εταιρείες που αναλαμβάνουν με τεχνητό ή και μέσω του ανθρώπινου παράγοντα να δημιουργήσουν μηνύματα spam. Το σχήμα που ακολουθεί

παρουσιάζει εύληπτα τις ροές εργασίας που έχουν δημιουργηθεί καθώς και τη σχέση των επιχειρήσεων σε αυτό το περιβάλλον. Βασική προϋπόθεση αυτής της προσέγγισης είναι η ύπαρξη μιας ομάδας εργαζομένων που είναι πρόθυμοι να επιλύσουν διαδραστικά ένα σύνολο CAPTCHAs σε αντάλλαγμα τη χρηματική τους αμοιβή.

Η παλαιότερη περιγραφή αυτού του συστήματος που έχει έχουμε βρει για μια τέτοια σχέση είναι σε ένα blog της εταιρείας Symantec τον Σεπτέμβριο 2006, η οποία τεκμηριώνει μέσω μιάς διαφήμιση για μία θέση πλήρους απασχόλησης που αφορά στην επίλυση CAPTCHA (Symantec). Συνήθως η αμοιβή για την επίλυση ενός CAPTCHA είναι 1 δολ. Από τότε που εμφανίστηκε η συγκεκριμένη διαφήμιση, η εμφάνιση παρόμοιων πολλαπλασιάστηκε. Εξειδικευμένες εταιρείες ιδρύθηκαν, όπως το Decaptcher προσφέροντας εξειδικευμένες υπηρεσίες.



Εικόνα 4-55 Το οικονομικό σύστημα επίλυσης

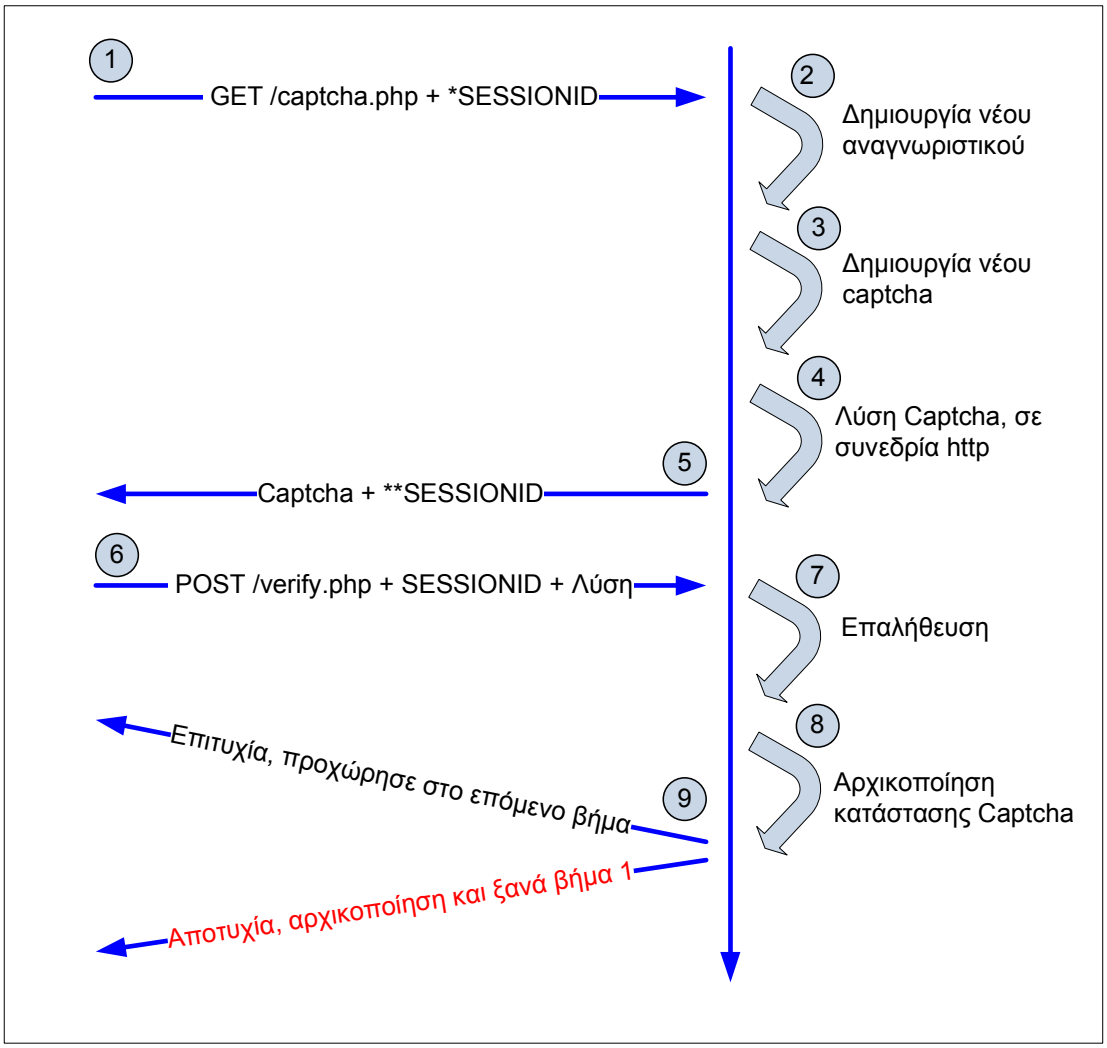
Όμως από τη στιγμή που αναπτύχθηκε η συγκεκριμένη βιομηχανία παρουσιάστηκε μία δραστική μείωση των ατόμων που απασχολούνται σε αυτή. Ταυτόχρονα μειώθηκε και η αμοιβή που προβλεπόταν για την ανθρώπινη εργασία. Από τη στιγμή που η συγκεκριμένη εργασία δεν απαιτεί εξειδικευμένο προσωπικό η παρεχόμενη αμοιβή λογικό ήταν να παρουσιάσει μείωση (Motoyama και συν., 2010). Το εργατικό δυναμικό που θα πλαισίωνε τις συγκεκριμένες εταιρείες αντλήθηκε από χώρες τριτοκοσμικές, ενώ δεν πρέπει να παραβλέψουμε και την πρόοδο που επήλθε στις μεθόδους αυτόματης επίλυσης.

Οι εταιρίες αυτές, προσφέρουν τις υπηρεσίες τους μέσω αμοιβής και παρέχουν τα αντίστοιχα APIs για την αυτοματοποίηση της διαδικασίας. Βέβαια η επίλυση γίνεται

χειροκίνητα, αλλά τόσο η παροχή του Captcha, όσο και η ανάκτηση της λύσης γίνεται αυτόματα.

1.20.3. BYPASS

Η υλοποίηση ενός Catpcha, για την παροχή προστασίας σε φόρμες περιλαμβάνει πεδία που αφορούν στοιχεία μιας δοσοληψίας.



Εικόνα 4-56 Δοσοληψία παροχής και επίλυσης captcha

Η δοσοληψία αυτή περιλαμβάνει τα εξής στάδια:

- ↓ Ο πελάτης απαιτεί ένα Captcha από τον εξυπηρετητή με ένα έγκυρο ή όχι αναγνωριστικό (session id).
- ↓ Αν ο πελάτης δεν προμηθεύσει ένα έγκυρο αναγνωριστικό, ένα νέο αναγνωριστικό παράγεται και η συνεδρία τεκμηριώνεται.
- ↓ Η λύση του Captcha αποθηκεύεται σε μία συνεδρία HTTP.

↓ Η εικόνα Captcha αποστέλλεται στον πελάτη. Αν ο πελάτης δεν απαντήσει με ένα έγκυρο αναγνωριστικό, επιστρέφεται το νέο αναγνωριστικό.

↓ Ο εξυπηρετητής ανακτά την λύση από την συνεδρία και την συγκρίνει με αυτή του πελάτη.

↓ Στη μεριά του εξυπηρετητή η λύση παρέχεται σε μορφή κειμένου.

↓ Αν η επαλήθευση είναι επιτυχής, ο πελάτης προχωρά στο επόμενο λογικό βήμα. Αν όχι ζητείται ξανά επιβεβαίωση.

Όπως φαίνεται από τη διαδικασία, η πιστοποίηση γίνεται στη μεριά του εξυπηρετητή. Πολλές όμως υλοποιήσεις χρησιμοποιούν τον πελάτη για αυτό. Κυρίως χρησιμοποιούνται κρυμμένα πεδία σε μια φόρμα για την αποστολή αυτής της πληροφορίας. Ένας επιτιθέμενος, μπορεί να επέμβει με δικές του τιμές και εφόσον η πιστοποίηση δεν γίνεται στον εξυπηρετητή δεν χρειάζεται να έχει πρόσβαση στην πρωτότυπη λύση.

1.20.3.1. Chosen Captcha Attack

Επίσης έχει διαπιστωθεί ότι πολλές υλοποιήσεις που βασίζονται σε κώδικα Javascript δεν χρησιμοποιούν τον εξυπηρετητή για επικύρωση (Gursev, 2012).

Η τεχνική που χρησιμοποιείται σε αυτή την περίπτωση ονομάζεται «chosen captcha text attack". Η διαδικασία που ακολουθείται κατά την επικύρωση είναι:

↓ Στην σελίδα της φόρμας, χρησιμοποιείται κώδικας JavaScript για την παραγωγή ενός τυχαίου αριθμού.

↓ Το τυχαίο νούμερο, αποστέλλεται μαζί με το αναγνωριστικό στον εξυπηρετητή για την παραγωγή της εικόνας captcha.

↓ Ο εξυπηρετητής παράγει την εικόνα βάσει του αριθμού που έλαβε. Ο αριθμός επίσης αποθηκεύεται στη συνεδρία http, για λόγους επαλήθευσης.

↓ Η εικόνα λαμβάνεται και εμφανίζεται στην φόρμα του πελάτη.

Ο επιτιθέμενος αντίστοιχα για να εκμεταλλευτεί την διαδικασία, μπορεί να:

↓ Προσκομίσει ένα έγκυρο αναγνωριστικό

↓ Θέσει την τιμή του Captcha στην συνεδρία http χρησιμοποιώντας το αναγνωριστικό.

↓ Να κάνει μία καταχώριση στη φόρμα με αυτή τη τιμή του Captcha στη συνεδρία και να παρακάμψει την ασφάλεια.

1.20.3.2. Captcha Rainbow Tables

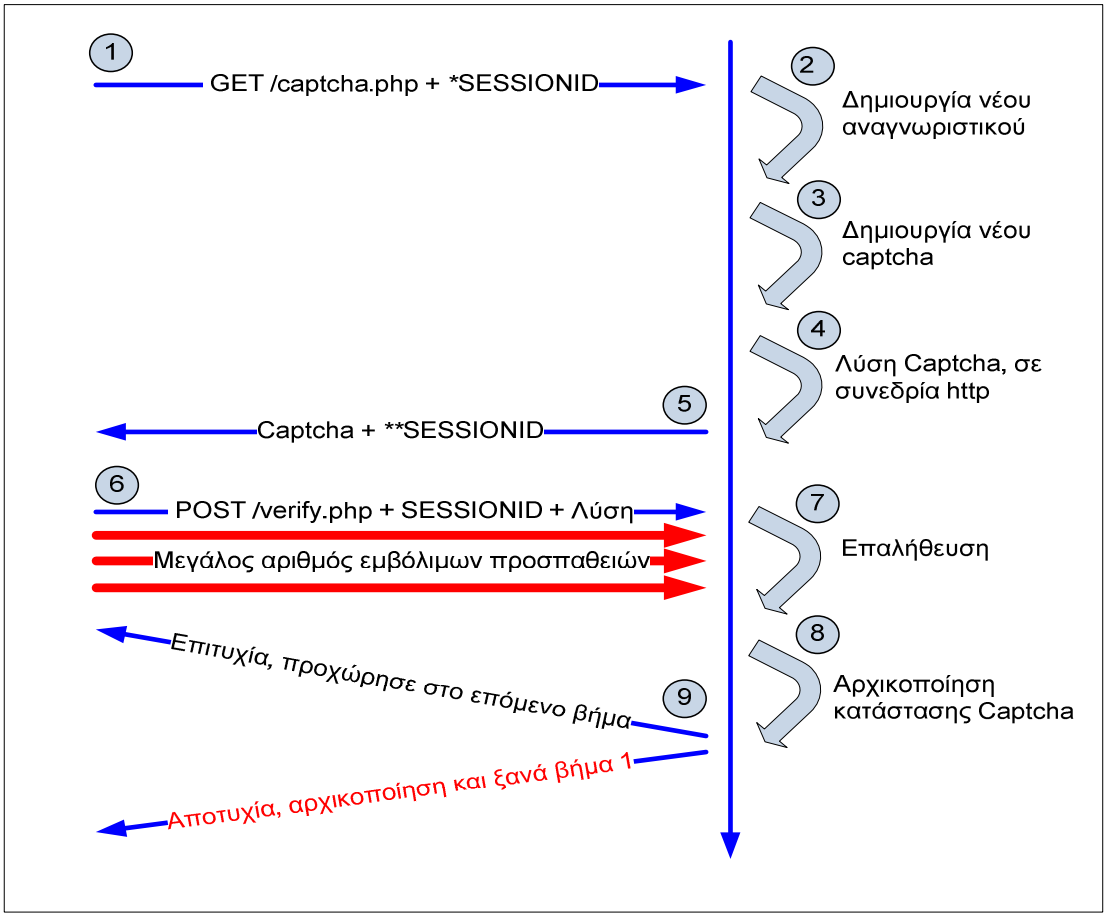
Μια άλλη τακτική που χρησιμοποιείται καλείται “captcha rainbow tables”. Πολλές υλοποιήσεις περιλαμβάνουν έναν πεπερασμένο αριθμό captcha, τα οποία επαναλαμβάνονται. Επειδή συνήθως αυτά αποθηκεύονται στο πίνακα μίας βάσης, περιέχουν αναγνωριστικά. Αυτά τα αναγνωριστικά αποστέλλονται μέσω κρυμμένων πεδίων σαν μέρος της διεύθυνσης κατά την ανάκτηση των Captcha. Επιπρόσθετα κάποιοι διαχειριστές ιστοτόπων δεν αλλάζουν ποτέ αυτά τα αναγνωριστικά. Σε αυτές τις περιπτώσεις τα Captcha μπορούν να ληφθούν μαζικά και να επιλυθούν τοπικά είτε μέσω OCR, είτε χειροκίνητα. Στη συνέχεια δημιουργείται ένας πίνακας με τη τιμή του αναγνωριστικού, καθώς και της λύσης. Συνεπώς αυτό οδηγεί στην αυτοματοποιημένη επίλυση για τον συγκεκριμένο ιστότοπο, αφού αρκεί μια απλή αναζήτηση στον πίνακα βάσει του αναγνωριστικού που διαβάζεται στο κρυμμένο πεδίο.

1.20.3.3. Captcha Brute-Forcing

Καλείται “In-Session CAPTCHA Brute-Forcing”, είναι ευρέως διαδεδομένη έχει τα εξής χαρακτηριστικά:

- Υποτίθεται ότι ο πελάτης ακολουθεί «ευλαβικά» τις οδηγίες του εξυπηρετητή για την έκδοση νέου Captcha, αν ο προηγούμενος έλεγχος απέτυχε.
- Ο κώδικας που δημιουργεί νέο Captcha και τον θέτει στην http συνεδρία, δουλεύει ανεξάρτητα του κώδικα που εκτελεί την επαλήθευση.
- Ο κώδικας που επαληθεύει δεν αφαιρεί την λύση από την συνεδρία και ως εκ τούτου επιτρέπει πολλαπλές απόπειρες επαλήθευσης στην ίδια περίοδο συνεδρίας.

Για την εκμετάλλευση αυτής της αδυναμίας, ένας επιτιθέμενος, όπως φαίνεται και στο σχήμα που ακολουθεί, μπορεί να κατευθύνει αρκετές προσπάθειες απ’ ευθείας στη διεύθυνση URL, που εκτελεί την επαλήθευση. Οι προσπάθειες και η αποστολή των Captcha, μπορεί να γίνει χρησιμοποιώντας ένα λεξικό με έτοιμους κωδικούς Captcha.



Εικόνα 4-57 In session brute force

1.21. ΠΡΟΓΡΑΜΜΑΤΑ ΠΑΡΑΚΑΜΨΗΣ

Υπάρχουν αρκετές υπηρεσίες Captcha, που έρχονται να βοηθήσουν τους δημιουργούς ιστοσελίδων. Χρησιμοποιώντας αυτές τις υπηρεσίες, μπορεί να μεγιστοποιηθεί η ασφάλεια ενός ιστοτόπου.

Στον αντίποδα, όπως είδαμε, υπάρχουν αντίστοιχες υπηρεσίες επίλυσης. Δραστηριότητες όπως «τοποθέτηση δικτυακού περιεχομένου», άρθρα, κριτικές, σχόλια, αυτόματα προφίλ, συμβαίνουν καθημερινά με σκοπό άλλοτε την διαφήμιση, ή την αλλοίωση της ποιότητας ενός ιστοτόπου. Αυτές οι διαδικασίες, μπορούν να επιτευχθούν αυτόματα με τη χρήση αντίστοιχου λογισμικού αντί – Captcha.

Αντίστοιχες εφαρμογές προγραμμάτων επίλυσης captcha, οι οποίες χρησιμοποιούν μηχανές αναγνώρισης ή και υπηρεσίες είναι:

- Xrumer

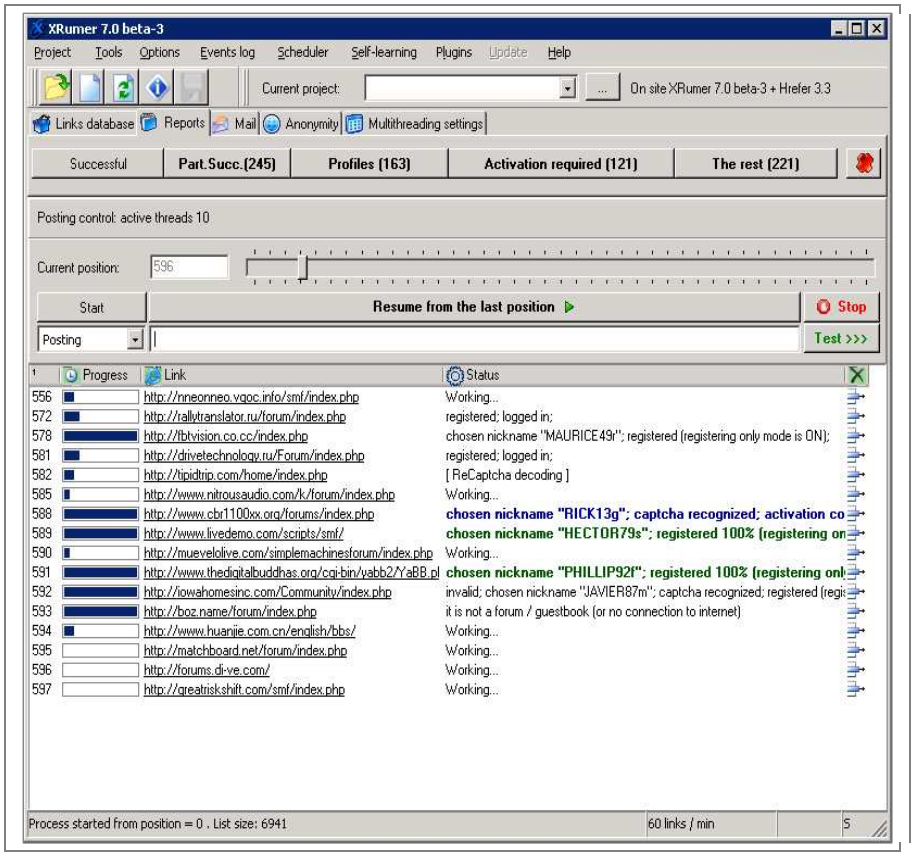
- GSACaptchaBraker
- Captcha Snipper
- reCaptchaOCR
- TesserCap

1.21.1. Xrumer

Το Xrumer, είναι ένα γνωστό εργαλείο, για την δημιουργία αυτόματων σχολίων σε φόρουμ. Χρησιμοποιεί ένα εξελιγμένο σύστημα παράκαμψης anti-spam τεχνικών, συμπεριλαμβανομένων και των Captcha. Πρωτοεμφανίστηκε το 2006 και είναι εμπορική εφαρμογή.

Χρησιμοποιεί μια υβριδική προσέγγιση επίλυσης, η οποία χρησιμοποιεί σε πρώτη φάση έναν αλγόριθμο για την αυτόματη επίλυση, ενώ αν αυτή αποτύχει σε δεύτερο επίπεδο ζητάει την βοήθεια του χρήστη για την εκπαίδευσή του.

Με την τελευταία έκδοση, σύμφωνα με τις προδιαγραφές του κατασκευαστή έχει τη δυνατότητα να επιλύει γραφικά Captcha ή κειμενικά.

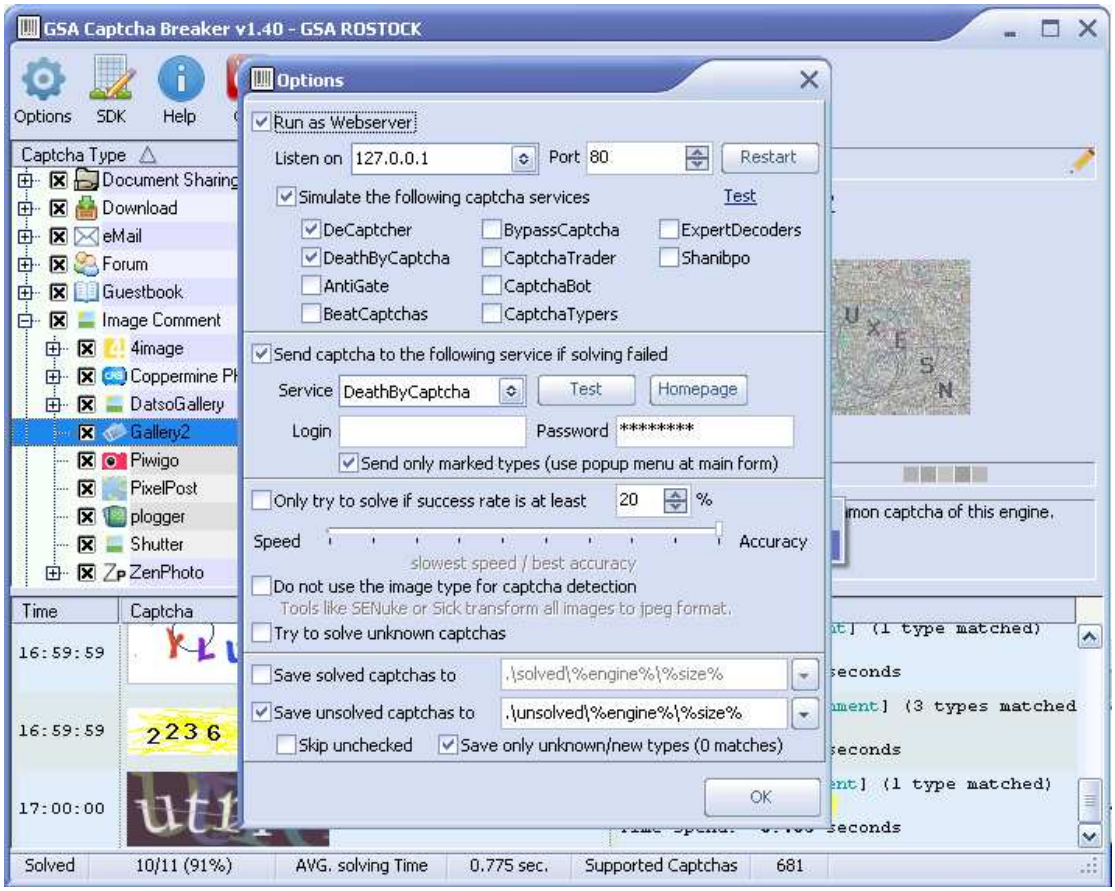


Εικόνα 4-58 XRumer

1.21.2. GSACaptchaBraker

Το GSACaptchaBraker, είναι και αυτό μια εμπορική εφαρμογή, για την επίλυση των πιο δημοφιλών Captcha. Αποτελείται από πολλαπλές μηχανές OCR, για μεγαλύτερη αποτελεσματικότητα.

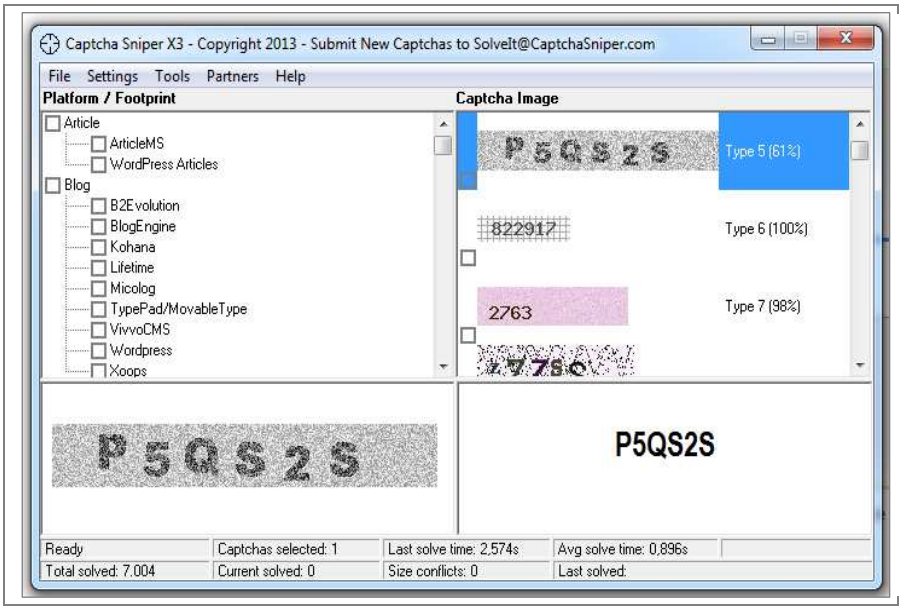
Μια δυνατότητα, είναι ότι μπορεί να στείλει τα άλυτα catpcha, αυτόματα σε υπηρεσίες επίλυσης και να επιστρέψει πάλι αυτόματα την λύση. Οι υπηρεσίες αυτές, είναι κυρίως συνδρομητικές και επιλύουν χειροκίνητα το captcha και επιστρέφουν άμεσα το αποτέλεσμα, εξομοιώνοντας ουσιαστικά την αυτόματη λύση. Μία τέτοια υπηρεσία για παράδειγμα είναι η Death By Captcha.



Εικόνα 4-59 GSA Captcha Breaker

1.21.3. Captcha Snipper

Το Captcha Snipper, δεν χρησιμοποιεί ενσωματωμένη μηχανή επίλυσης, αλλά κάνει κλήσεις σε υπηρεσίες επίλυσης όπως: Decaptcha, Death By Captcha, AntiCaptcha και ByPass Captcha.



Εικόνα 4-60 Captcha Snipper

1.21.4. reCaptchaOCR

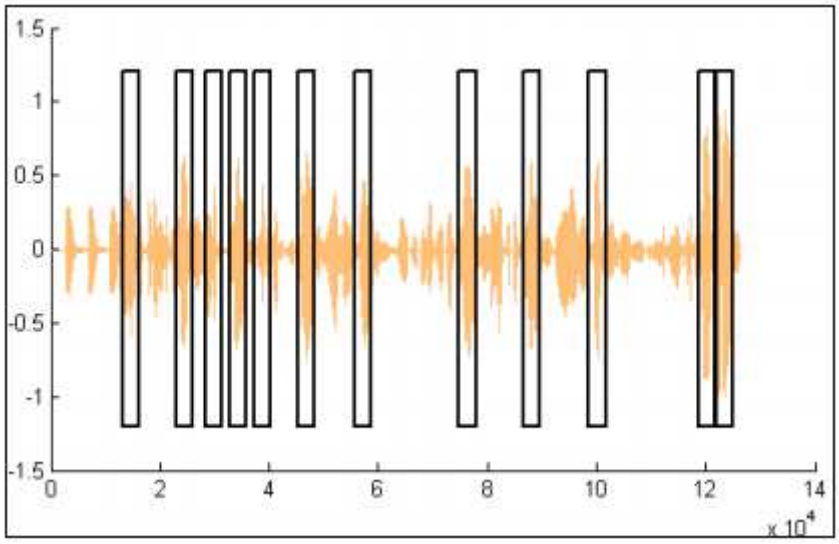
Το reCaptchaOCR, ειδικεύεται στην επίλυση του reCaptcha το οποίο διατίθεται από την Google. Ένα πρόσθετο χαρακτηριστικό του, είναι η δυνατότητα ενσωμάτωσης σαν πρόσθετο σε άλλα προγράμματα. Αυτό επιτυγχάνεται, με την αγορά της αντίστοιχης DLL βιβλιοθήκης ή απ' ευθείας κώδικα παρεμβολής.

Ουσιαστικά αποτελεί μια μηχανή, για την δημιουργία εφαρμογών anti – captcha.

1.21.5. Παράκαμψη Captcha ήχου

Η παράκαμψη ενός ηχητικού captcha, δεν υιοθετείται από κάποιο πρόγραμμα. Αντιθέτως έχει παρουσιαστεί μία τεχνική από τους Jennifer Tam, Jiri Simsa, Sean Hyde & Luis Von Ahn.

Η προσέγγιση για την παράκαμψη ξεκινά πρώτα με τον διαχωρισμό του αρχείου ήχου σε τμήματα θορύβου, που ανταποκρίνονται στις λέξεις. Στη συνέχεια αυτά τα τμήματα χαρακτηρίζονται από ετικέτες σήμανσης. Τα τμήματα με την υψηλότερη συχνότητα ταξινομούνται χρησιμοποιώντας τεχνικές εκμάθησης. Η εκμάθηση γίνεται χειροκίνητα, για να ταυτοποιηθούν τα αντίστοιχα τμήματα σε ψηφία ή γράμματα.



Εικόνα 4-61 Τμηματοποίηση υψηλών συχνοτήτων

Αυτά τα εξαγόμενα τμήματα, που αντιστοιχούν σε χαρακτήρες ή ψηφία χρησιμοποιούνται στη συνέχεια για αντιπαραβολή με άλλα αρχεία ήχου για να γίνει η παράκαμψη.

Η τεχνική έχει παρουσιάσει σημαντική επιτυχία με αρκετούς πάροχους ηχητικών Captcha και τα αποτελέσματα της μεθόδου παρουσιάζονται αναλυτικά στην ανωτέρω μελέτη.

1.22. TesserCap

Το tesserCap είναι μια εφαρμογή ανοιχτού κώδικα, επίλυσης CATPCHA. Είναι η εφαρμογή η οποία θα χρησιμοποιήσουμε για τον έλεγχο παράκαμψης. Για αυτό τον λόγο θα προσπαθήσουμε να αναλύσουμε τα τεχνικά χαρακτηριστικά της, καθώς και τον τρόπο λειτουργίας της.

1.22.1. Χαρακτηριστικά TesserCap

Το TesserCap ουσιαστικά είναι ένα εργαλείο ανάλυσης το οποίο χρησιμοποιεί σαν μηχανή οπτικής αναγνώρισης την Tesseract για την επεξεργασία των Captcha. Μερικά από τα κύρια χαρακτηριστικά της είναι:

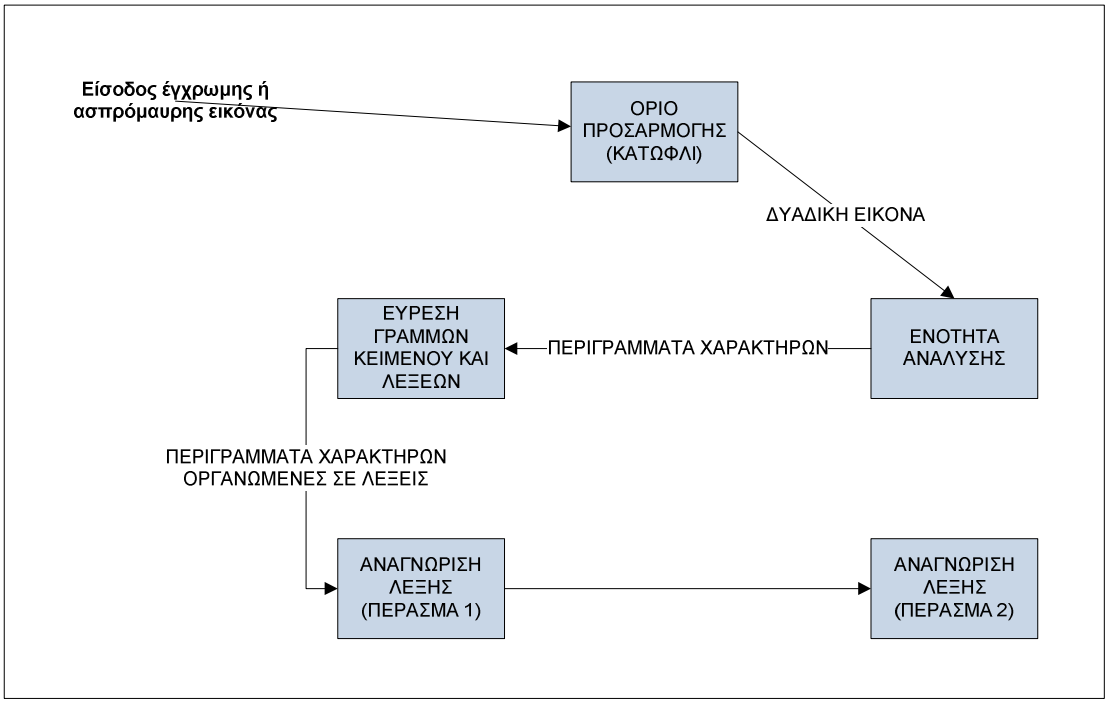
- **Υποστήριξη διαμεσολαβητή**: Ο διαμεσολαβητής χρησιμοποιείται κυρίως για την απόκρυψης της IP διεύθυνσής μας κατά τον έλεγχο. Αυτή επίσης μπορεί να αλλάζει δυναμικά, από μία λίστα proxies.

- **Στατιστική ανάλυση**: Πρόκειται για ενότητα εμφάνισης στατιστικών στοιχείων που αφορούν την αποτελεσματικότητα του προγράμματος.
- **Σύνολα χαρακτήρων**: Δυνατότητα επιλογής συγκεκριμένης ομάδας χαρακτήρων κατά τον έλεγχο.

1.22.2. Μηχανή οπτικής αναγνώρισης Tessercat

Η μηχανή οπτικής αναγνώρισης Tessercat έχει αναπτύχθηκε από την HP-UX έως το 1994. Από το 1995 αποτελεί και αυτή λογισμικό ανοιχτού κώδικα και υποστηρίζεται από την Google.

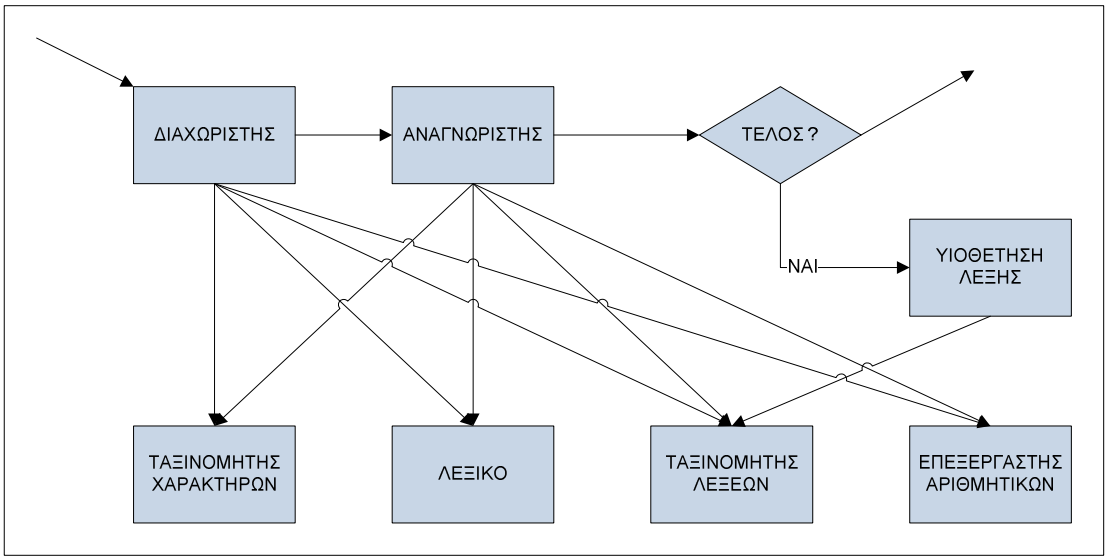
Η αρχιτεκτονική της μηχανής παρουσιάζεται στο διάγραμμα που ακολουθεί:



Εικόνα 4-62 Στάδια επεξεργασίας Tesseract

Η διαδικασία που ακολουθείται είναι σειριακή και αποτελείται από διαδοχικά βήματα. Το πρώτο βήμα περιλαμβάνει την εισαγωγή της εικόνας, ενώ μετά ακολουθεί η ανάλυση της και η εξαγωγή των περιγραμμάτων σύμφωνα με ένα δοσμένο κατώφλι. Τα περιγράμματα αυτά, χωρίζονται σε τμήματα, τα οποία με τη σειρά τους οργανώνονται σε γραμμές κειμένου. Οι γραμμές κειμένου αναλύονται ως

προς την προοπτική τους και στη συνέχεια χωρίζονται σε διαφορετικές λέξεις ανάλογα με την μεταξύ τους απόσταση[2].



Εικόνα 4-63 Αναγνώριση λέξης

Η αναγνώριση του κειμένου είναι διαδικασία διπλής επεξεργασίας. Κατά την πρώτη, γίνεται μια προσπάθεια αναγνώρισης από τον λεκτικό επεξεργαστή. Στη συνέχεια γίνεται και δεύτερο πέρασμα, χρησιμοποιώντας τις λέξεις οι οποίες έχουν ανακαλυφθεί κατά το πρώτο πέρασμα για την εκπαίδευση του αναλυτή.

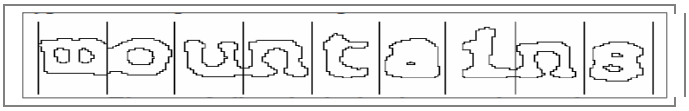
Κατά το πρώτο στάδιο ανάλυσης υπολογίζονται οι αποστάσεις των γραμμών. Αυτές καθορίζονται από συντεταγμένες καλύπτοντας έτσι και τις περιπτώσεις που μια γραμμή δεν είναι σαρωμένη οριζόντια (Rousseeuw & Leroy, 2003).

Μόλις βρεθούν οι γραμμές του κειμένου, υπολογίζονται οι γραμμές βάσεις με μεγαλύτερη ακρίβεια. Αυτό είναι και το κύριο πλεονέκτημα της μηχανής Tesseract, επειδή επιτρέπει την αναγνώριση ακόμα και καμπύλου κειμένου.

Εικόνα 4-64 Εύρεση γραμμών βάσης

Η εύρεση του διαστήματος ανάμεσα στους χαρακτήρες είναι σχετικά απλή διαδικασία αν το κείμενο έχει σταθερά διαστήματα. Αν το tesseract βρει ένα σταθερό διάστημα τεμαχίζει την λέξη σύμφωνα με αυτό.

[2] Μεγαλύτερη απόσταση, σημαίνει ένα κενό διάστημα μεταξύ δύο λέξεων.

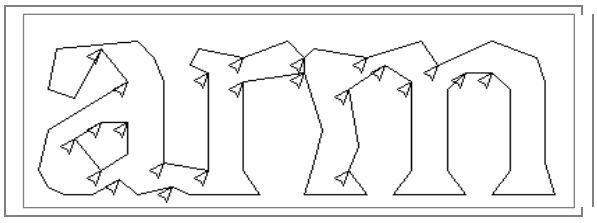


Εικόνα 4-65 Κείμενο με σταθερό διάστημα

Στην περίπτωση όμως που το διάστημα δεν είναι σταθερό, τα πράγματα περιπλέκονται και το tesseract λύνει αυτά τα προβλήματα με την μέτρηση διαστημάτων ανά μικρότερες περιοχές και η τελική απόφαση λαμβάνεται με την μέθοδο της σύγκρισης σχεδίων (χαρακτήρων).

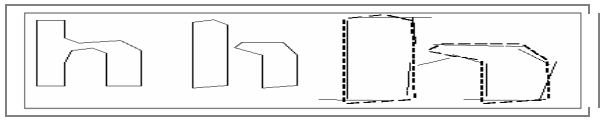
Εικόνα 4-66 Κείμενο χωρίς σταθερό διάστημα ανάμεσα στους χαρακτήρες

Το στάδιο αναγνώρισης των χαρακτήρων, περιλαμβάνει τις απαραίτητες διαδικασίες για την εξαγωγή τους από τη λέξη. Υπάρχουν περιπτώσεις όπου οι χαρακτήρες επικαλύπτονται. Το tesseract, για την βελτίωση των αποτελεσμάτων προσπαθεί να διαχωρίσει τους χαρακτήρες, ανιχνεύοντας σημεία από κοίλες κορυφές χρησιμοποιώντας πολυγωνική προσέγγιση. Μπορεί να χρειαστούν έως και τρία ζεύγη σημείων για να είναι ο διαχωρισμός επιτυχής.



Εικόνα 4-67 Διαχωρισμός ενωμένων χαρακτήρων

Ένα άλλο πρόβλημα που μπορεί να προκύψει κατά την εξαγωγή των χαρακτήρων, είναι η αναγνώριση «σπασμένων» σχεδίων.



Εικόνα 4-68 Αναγνώριση σπασμένων χαρακτήρων

Αυτό το πρόβλημα μπορεί να προκύψει είτε απευθείας από το σαρωμένο κείμενο, είτε από την προηγούμενη διαδικασία του διαχωρισμού. Στη συνέχεια ο

αναγνωριστής, με τη χρήση γράφων και διαδοχικών περασμάτων, αναλαμβάνει την ταυτοποίηση των χαρακτήρων.

1.23. ΕΛΕΓΧΟΣ ΠΑΡΑΚΑΜΨΗΣ

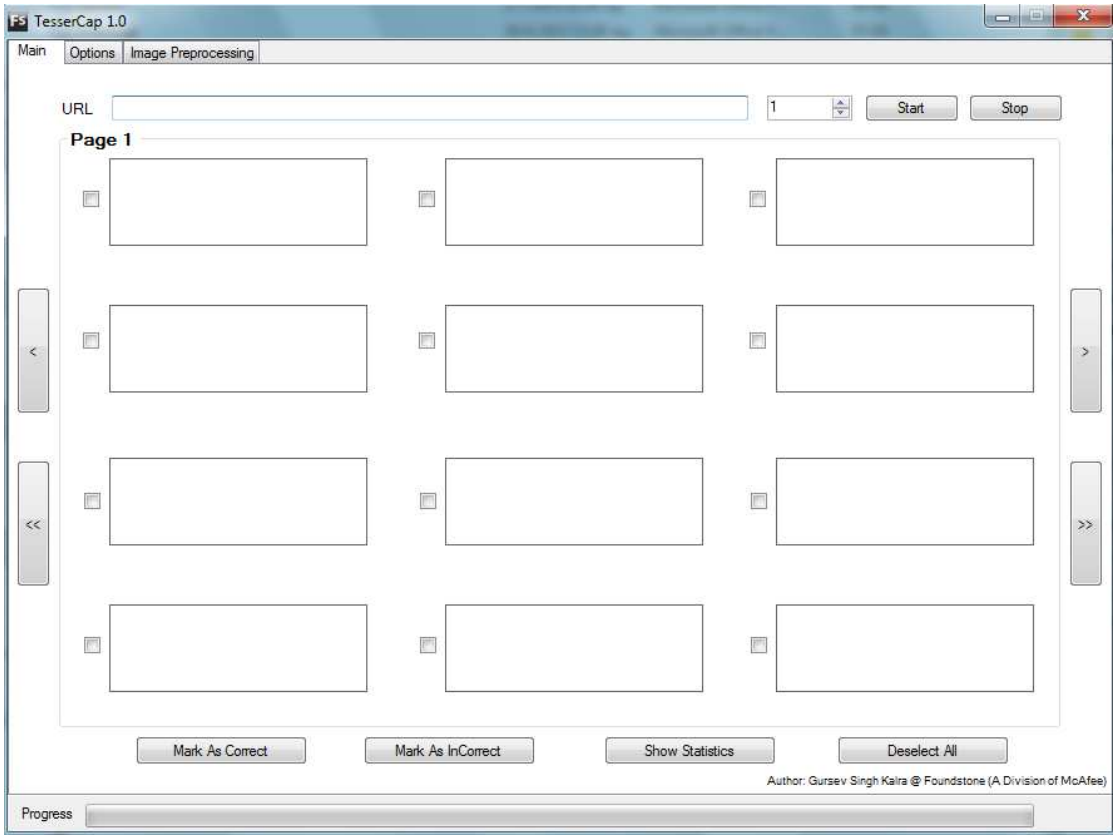
Ό έλεγχος της παράκαμψης, όπως έχουμε ήδη αναφέρει θα γίνει με τη χρήση του προγράμματος tesserCap. Η εγκατάσταση του προγράμματος είναι απλή, αφού γίνεται μέσω οδηγού. Το tesserCap είναι ένα εργαλείο με γραφική διεπιφάνεια. Έτσι μετά την εκτέλεση του προγράμματος ανοίγει το κύριο παράθυρο της εφαρμογής.

1.24. ΠΕΡΙΓΡΑΦΗ ΠΡΟΓΡΑΜΜΑΤΟΣ

Στο παράθυρο αυτό βάζουμε την διεύθυνση URL, η οποία περιέχει την εικόνα Captcha. Θα πρέπει να διευκρινίσουμε ότι η διεύθυνση αυτή αντιστοιχεί στην πραγματική διεύθυνση η οποία παράγει την εικόνα και όχι την διεύθυνση URL του ιστοτόπου. Αυτό είναι ένα μειονέκτημα του συγκεκριμένου προγράμματος, διότι ο χρήστης θα πρέπει να την εισάγει χειροκίνητα, καθώς αυτή δεν μπορεί να βρεθεί αυτόματα από το πρόγραμμα. Η τακτική που θα ακολουθήσουμε είναι να σύρουμε την εικόνα captcha στο παράθυρο του φυλλομετρητή για να επιστρέψει η πραγματική διεύθυνση, για να μπει στο συγκεκριμένο πεδίο του TesserCap.

Το διπλανό πεδίο αφορά τον αριθμό των προσπαθειών, ενώ η επόμενη ενότητα εμφανίζει τα αποτελέσματα της επεξεργασίας.

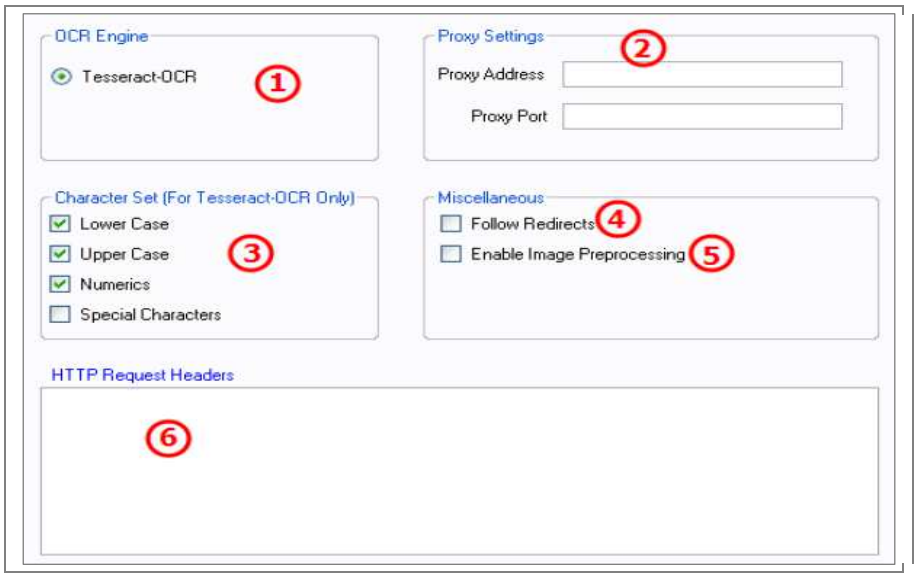
Με τα κουμπιά που βρίσκονται στο τέλος της σελίδας, μπορούμε να επιλέξουμε ποια από τα αποτελέσματα της επεξεργασίας ήταν σωστά και να λάβουμε τα αντίστοιχα στατιστικά στοιχεία.



Εικόνα 4-69 Αρχική οθόνη TesserCap

Η επόμενη καρτέλα, αφορά τις ρυθμίσεις της επεξεργασίας και οι οποίες είναι:

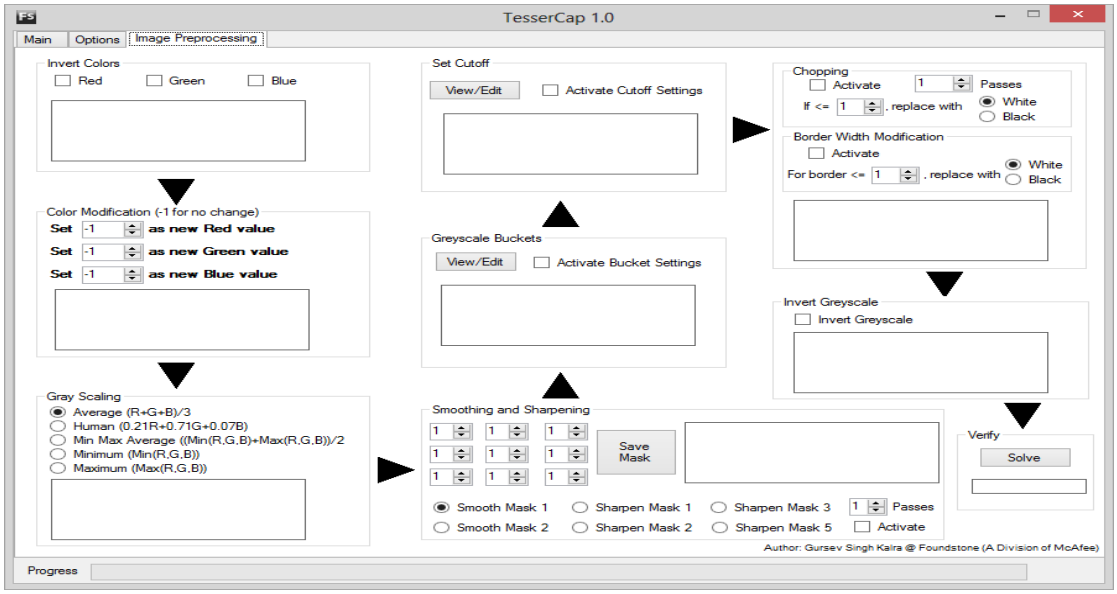
- Μηχανή: Η μηχανή OCR που χρησιμοποιείται. Στην περίπτωσή μας υποστηρίζεται μόνο η Tesseract.
- Σύνολο χαρακτήρων: Ρυθμίζουμε το σετ χαρακτήρων για την αναγνώριση. Στην περίπτωση για παράδειγμα που έχουμε μόνο αριθμούς πατάμε την αντίστοιχη επιλογή. Αυτό βοηθάει στην επιτάχυνση της επεξεργασίας και την βελτίωση των αποτελεσμάτων αν το Captcha περιέχει μόνο αριθμούς ή κάποιο άλλο σετ χαρακτήρων από τα προσφερόμενα.
- Ρυθμίσεις διαμεσολάβησης: Αυτό έχει να κάνει, με την απόκρυψη της πραγματικής μας διεύθυνσης με τη χρήση κάποιου proxy.
- Άλλες ρυθμίσεις: Στην περίπτωση που η εικόνα captcha, παράγεται μέσω μιας άλλης διεύθυνσης τότε θα πρέπει να ενεργοποιήσουμε την επιλογή “Follow Redirects”. Επίσης μπορούμε να δούμε, αλλά και να επέμβουμε στα στάδια επεξεργασίας, με την επιλογή “Enable Image Preprocessing”.



Εικόνα 4-70 Ρυθμίσεις TesserCap

Η τελευταία καρτέλα αφορά, την προ-επεξεργασία της εικόνας. Από αυτή ρυθμίζουμε διάφορες παραμέτρους όπως:

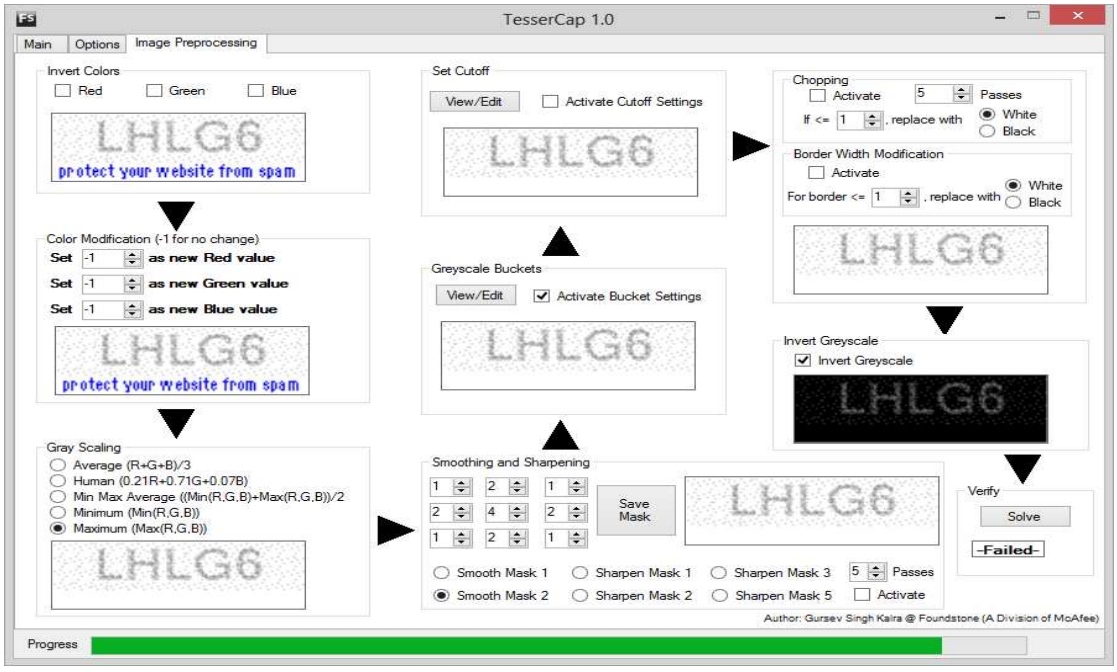
- **Αντιστροφή χρωμάτων**: Εδώ μπορούμε να ρυθμίσουμε τα βασικά χρώματα (RGB), και αυτό είναι ιδιαίτερα χρήσιμο σε μονοχρωματικές εικόνες που βασίζονται σε ένα από αυτά.
- **Ρύθμιση του κατωφλίου:** Μπορούμε να περιορίσουμε τον χώρο σάρωσης. Αυτό χρησιμοποιείται αν για παράδειγμα το Captcha έχει οριζόντιες ή κάθετες γραμμές εκτός του παραθύρου των χαρακτήρων που μπορούν να μπερδέψουν το πρόγραμμα.
- **Τεμαχισμός:** Ενεργοποιεί τον αυτόματο τεμαχισμό των ενωμένων χαρακτήρων
- **Ρύθμιση του γκρι:** Πως θα υπολογιστούν οι διαβαθμίσεις του γκρι.
- **Φιλτράρισμα της εικόνας:** Ο τρόπος φιλτραρίσματος της εικόνας. Περιέχονται δύο επιλογές (Smooth, Sharp) με τις αντίστοιχες παραλλαγές.



Εικόνα 4-71 Προ - Επεξεργασία εικόνας

1.25. ΦΟΡΜΑ ΕΠΙΚΟΙΝΩΝΙΑΣ

Στη φόρμα επικοινωνίας χρησιμοποιήσαμε, σαν πάροχο Captcha την snaphost. Οι αντίστοιχες ρυθμίσεις, με τις οποίες πετύχαμε το καλύτερο αποτέλεσμα παρουσιάζονται στην εικόνα που ακολουθεί.



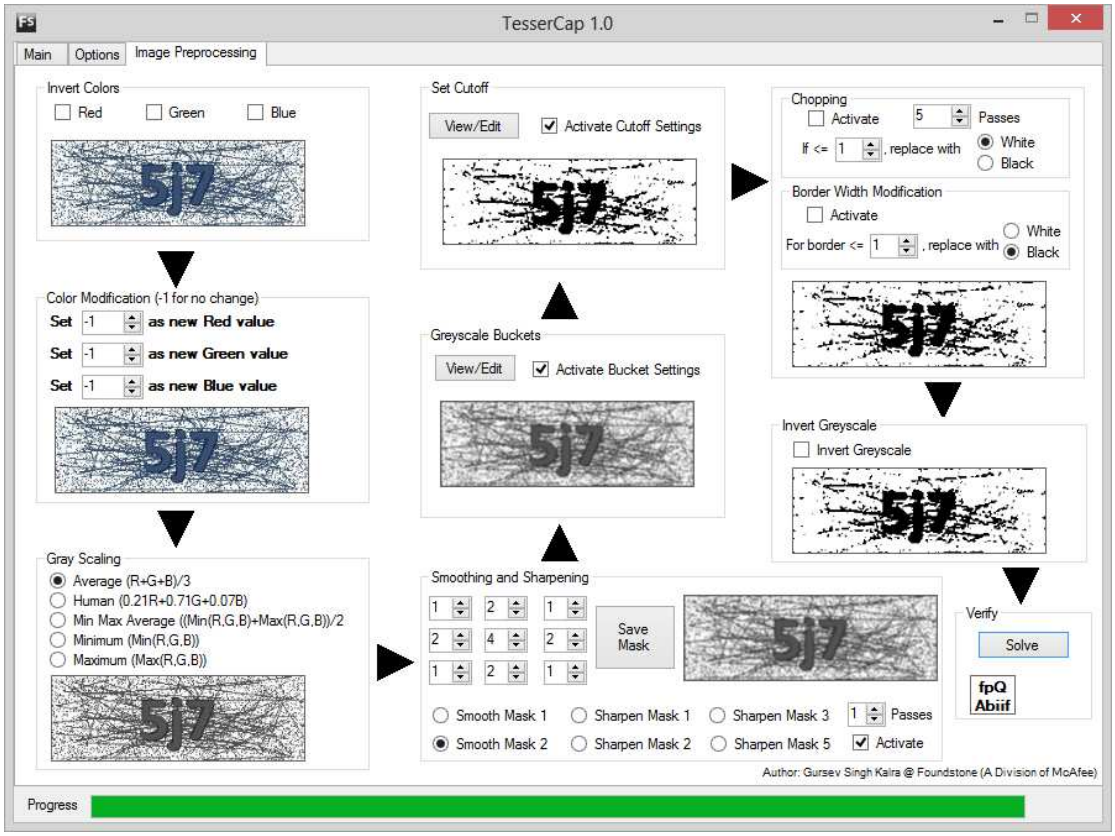
Εικόνα 4-72 Ρυθμίσεις SnapHost

Αντίστοιχα τα αποτελέσματα που λάβαμε ήταν 21 σωστές ανιχνεύσεις στις 100 προσπάθειες.

Εικόνα 4-73 Αποτελέσματα SnapHost

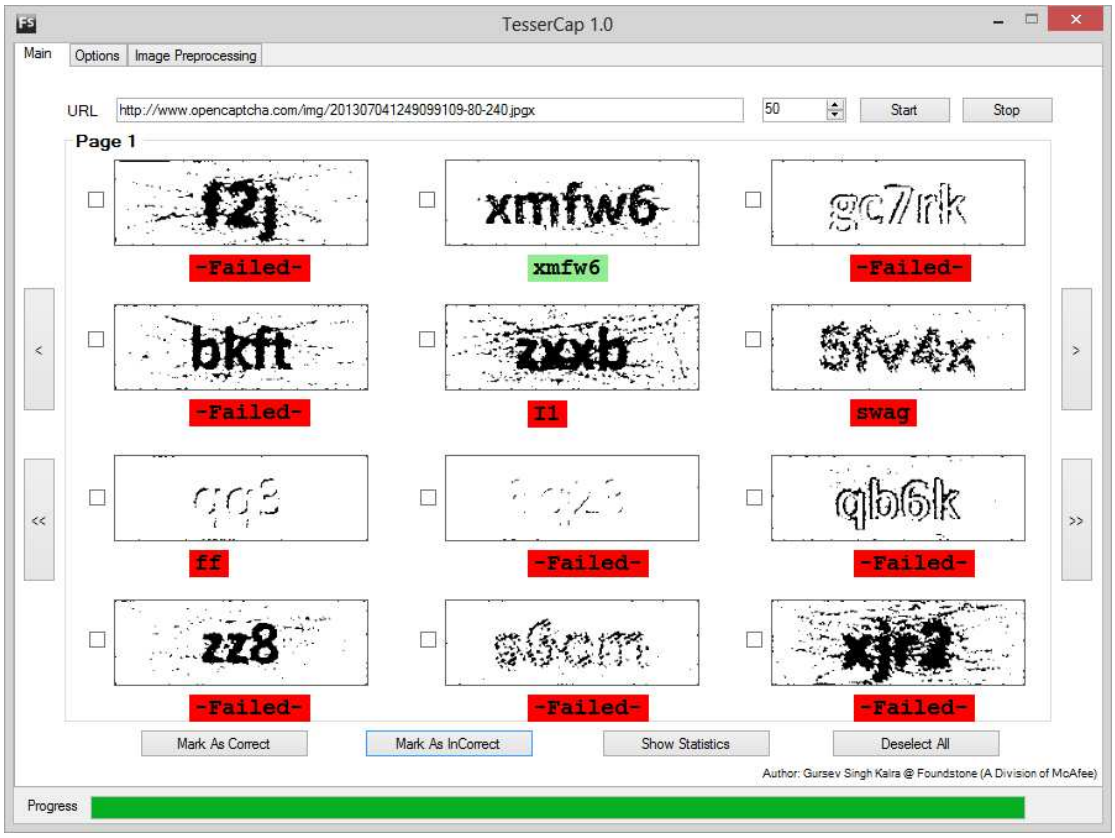
1.26. ΦΟΡΜΑ ΣΥΝΔΕΣΗΣ

Στην φόρμα που χρησιμοποιεί ο ιστότοπός μας, για την σύνδεση ενός χρήστη στο σύστημα χρησιμοποιήσαμε το opencaptcha.



Εικόνα 4-74 Ρυθμίσεις opencaptcha

Τα αντίστοιχα αποτελέσματα που λάβαμε ήταν 4 σωστές ανιχνεύσεις στις 100 προσπάθειες.

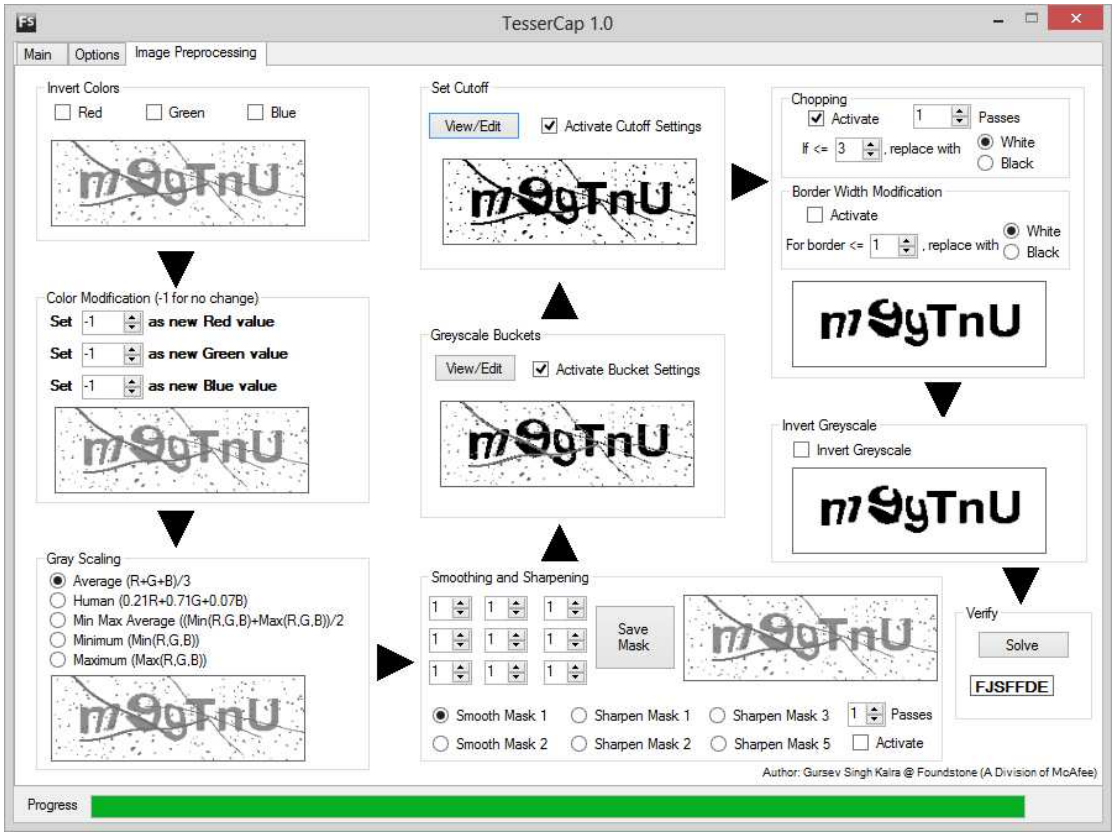


Εικόνα 4-75 Αποτελέσματα OpenCaptcha

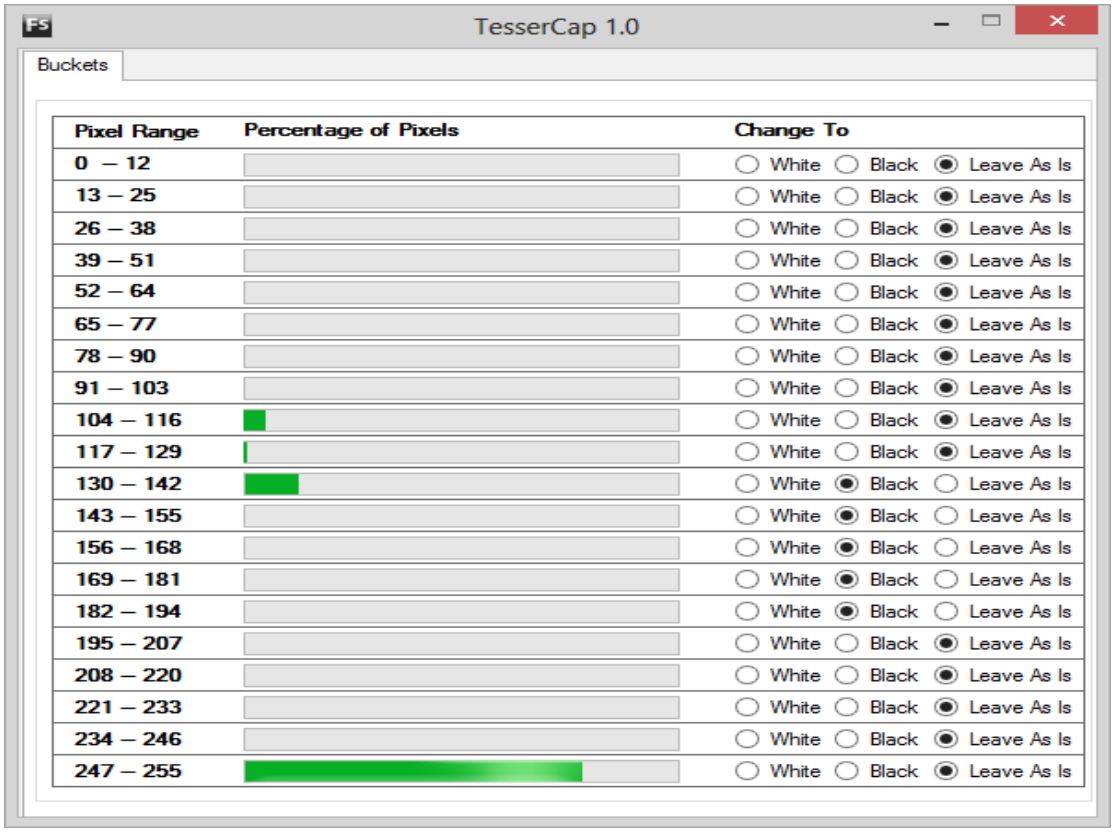
1.27. ΕΓΓΡΑΦΗ ΧΡΗΣΤΗ

Στη φόρμα της εγγραφής, χρησιμοποιήθηκε το secureimage, ως έλεγχος captcha. Οι ρυθμίσεις για να πάρουμε 16 επιτυχείς ανιχνεύσεις στις 100 προσπάθειες, ήταν αρκετά χρονοβόρες αφού χρειάστηκε να πειράξουμε και τις παραμέτρους cutoff και greyscale buckets.

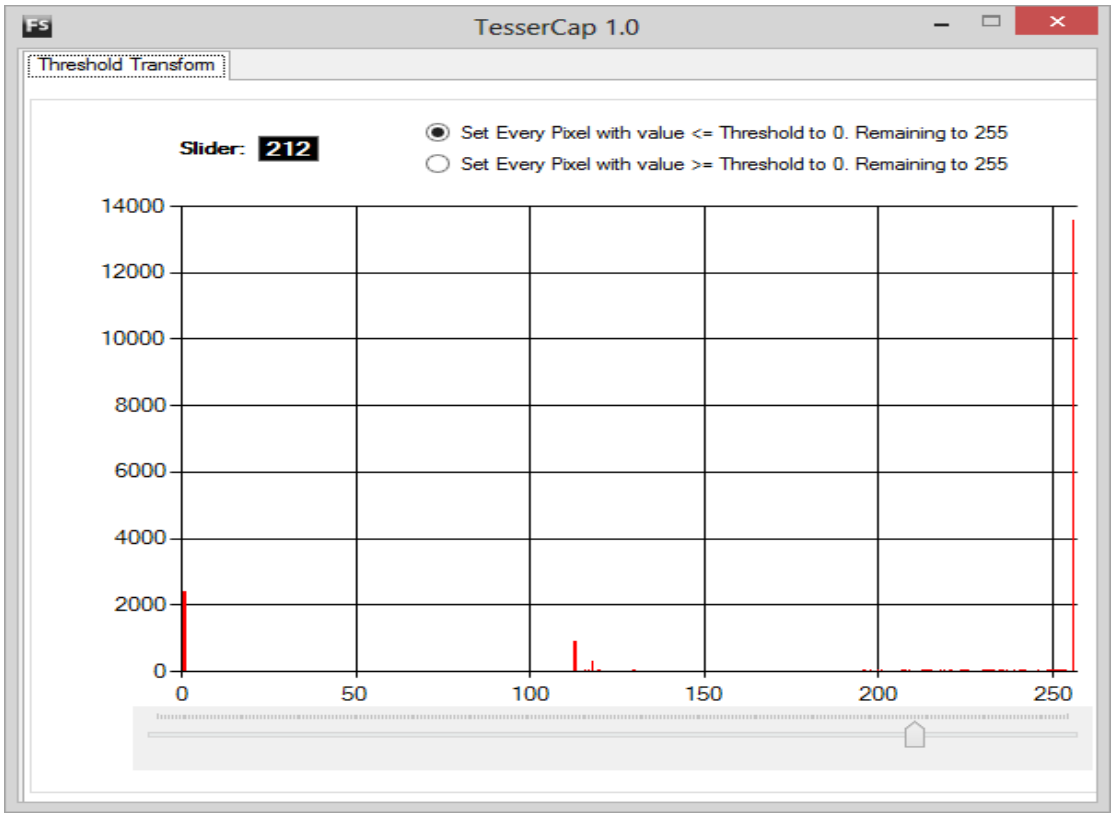
Οι ρυθμίσεις παρουσιάζονται στις εικόνες που ακολουθούν.



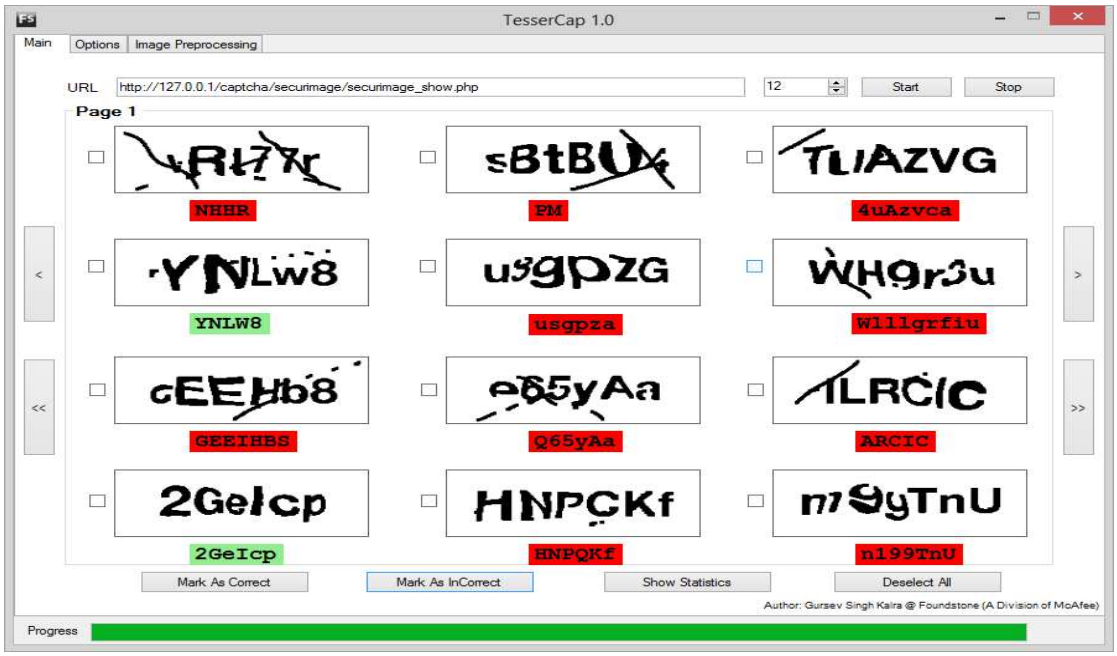
Εικόνα 4-76 Ρυθμίσεις για secureimage



Εικόνα 4-77 Ρυθμίσεις bucket για secureimage



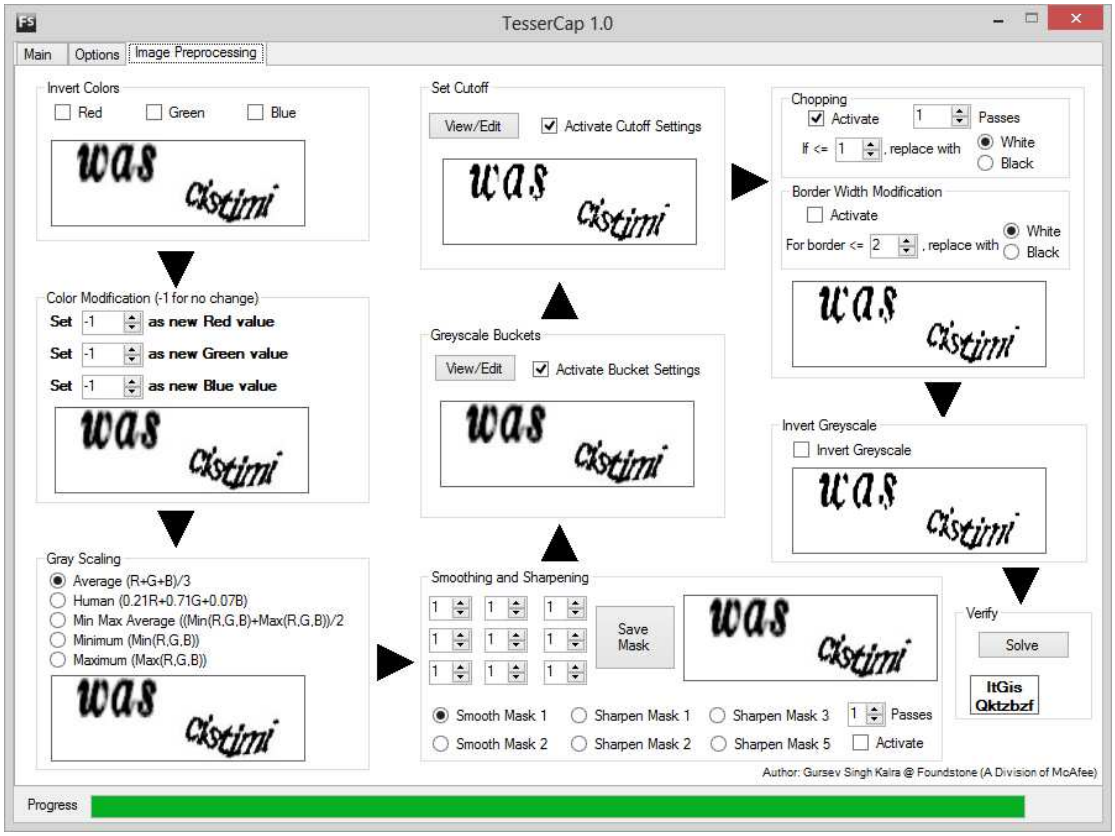
Εικόνα 4-78 Ρυθμίσεις cutoff



Εικόνα 4-79 - Αποτελέσματα για secureimage

1.28. ΣΧΟΛΙΑΣΜΟΣ ΚΑΤΑΣΤΗΜΑΤΩΝ

Τα σχόλια των χρηστών για κάθε κατάστημα, επικυρώνονται όπως είδαμε μέσω του reCaptcha. Στις δοκιμές μας με το TesserCap, ήταν το μόνο Captcha που δεν καταφέραμε να επιλύσουμε. Η δυσκολία ήταν κυρίως στη δεύτερη λέξη που χρησιμοποιείται, ενώ σε κάποιες περιπτώσεις η πρώτη ανακαλύφθηκε σωστά.



Εικόνα 4-80 Προσπάθεια επίλυσης reCaptcha

ΠΗΓΕΣ ΚΕΦΑΛΑΙΟΥ

- E. Bursztein, S. Bethard, J. C. Mitchell, D. Jurafsky, and C. Fabry. How good are humans at solving CAPTCHAs? a large scale evaluation. In IEEE S&P ’10, 2010.

-] J. Yan and A. S. El Ahmad. Usability of CAPTCHAs or usability issues in CAPTCHA design. In SOUPS ’08, pages 44–52, New York, NY, USA, 2008. ACM

- M. Motoyama, K. Levchenko, C. Kanich, D. McCoy, G.M. Voelker, and S. Savage. Re: CAPTCHAs-Understanding CAPTCHA-solving services in an economic context. In Proc. of the 19th USENIX Security Symposium, USENIX Association, 435—462

- P.J. Rousseeuw, A.M. Leroy, Robust Regression and Outlier Detection, Wiley-IEEE, 2003.

- K.S. Gursev Attacking CAPTCHAs for Fun and Profit, 2012

- T. Jennifer, J. Simsa, S.Hyde, L.V.Ahn Breaking Audio CAPTCHAs, 2008

ΣΥΜΠΕΡΑΣΜΑΤΑ

Το διαδίκτυο δεν είναι αποκλειστικά χώρος ενημέρωσης, αλλά και χώρος οικονομικής δραστηριότητας. Χωρίς να γίνεται αντιληπτό από την πλειοψηφία των χρηστών του, πίσω από το προσκήνιο μία μεγάλη βιομηχανία προσπαθεί με κάθε τρόπο να σπάσει τους προστατευτικούς μηχανισμούς των ιστοσελίδων και να περάσει το μήνυμά της. Η βιομηχανία αυτή είναι εξαιρετικά επικερδής και βασίζεται τόσο στον ανθρώπινο παράγοντα όσο και σε εξειδικευμένες τεχνικές μηχανικής μάθησης. Μέσα από τις τεχνικές αυτές ένα λογισμικό μπορεί να εισχωρήσει μαζικά και να περάσει το επιθυμητό μήνυμα.

Από την πλευρά της ιστοσελίδας η τεχνολογία CAPTCHA αναλαμβάνει να διασφαλίσει σε μεγάλο βαθμό το συγκεκριμένο φαινόμενο. Μπορεί για την ώρα να θεωρείται μία αποτελεσματική τεχνική, αλλά δεν αποτελεί πανάκεια. Πολλές φορές ζητήματα ευχρηστίας έρχονται στην επιφάνεια. Η ευχρηστία αποτελεί το τίμημα που πρέπει να πληρώσει η ιστοσελίδα για την ασφάλειά της. Η αφιέρωση μερικών δευτερολέπτων για την αντιγραφή των αριθμών ή των γραμμάτων που παρουσιάζει το CAPTCHA. Γενικά η διαδικασία αυτή μοιάζει να είναι ανούσια για τον χρήστη, αλλά σίγουρα δεν είναι για τους ανθρώπους που βρίσκονται πίσω από την εφαρμογή.

Η παρούσα εργασία προσπάθησε να κινηθεί μέσα σε διάφορα επίπεδα προκειμένου να καταστήσει κατανοητή τη λογική και την τεχνολογία του μηχανισμού. Αφενός επιχειρήθηκε να αναλυθεί η αναγκαιότητα του μηχανισμού CAPTCHA, τα κίνητρα που ώθησαν στην ανάπτυξη του και τις υπάρχουσες λύσεις που έχει δημιουργήσει η τεχνολογία.

Ταυτόχρονα στην παρούσα διπλωματική εργασία έγινε προσπάθεια να αναπτυχθεί λογισμικό που θα μπορούσε να ξεπεράσει σε ένα μεγάλο βαθμό τους περιορισμούς που θέτει η τεχνολογία CAPTCHA. Τα αποτελέσματα ήταν ενθαρρυντικά αποδεικνύοντας ότι ακόμα και η συγκεκριμένη τεχνολογία δεν μπορεί να θεωρηθεί ως άτρωτη. Ανεξάρτητα όμως από τα αποτελέσματα της παρούσας έρευνας η τεχνολογία θεωρείται ότι καλύπτει ικανοποιητικά τις ανάγκες του σημερινού χρήστη που θέλει να ασφαλίσει την ιστοσελίδα του. Η διαδικασία δημιουργίας ενός απαιτητικού λογισμικού αποδείχθηκε μία αρκετά απαιτητική διαδικασία που βασίζεται πάνω στη γνώση των ιδιοτήτων της τεχνολογίας CAPTCHA. Η καλή

γνώση της τεχνολογίας μπορεί να οδηγήσει στην αναγνώριση των τρωτών σημείων της, τα οποία θα πρέπει να κατασταθούν εκμεταλλεύσιμα.

Είναι προφανές ότι οι δείκτες αναγνώρισης μεταξύ ανθρώπου και μηχανής χωρίζονται από ένα βαθύ χάσμα. Η έρευνα θα πρέπει να συνεχίσει και να μειώσει αυτό το κενό μεταξύ αυτής της δυνατότητας. Τη διαδικασία αυτή θα πρέπει να την εκμεταλλευτεί η βιομηχανία που ασχολείται με τη διασφάλιση των ιστοσελίδων προκειμένου να αυξήσει την αποτελεσματικότητα τους. Είναι δεδομένο ότι δεν θα μπορούν οι τεχνολόγοι που αναπτύσσουν το CAPTCHA να το δυσκολεύουν επ' αόριστον. Οι δυνατότητες του ανθρώπου είναι περιορισμένες. Από κάποιο σημείο και μετά η δυσκολία θα καταστήσει την τεχνολογία δύσκαμπτη για τις ανάγκες του χρήστη. Ο ανθρώπινος παράγοντας πρέπει να λαμβάνεται σοβαρά υπ' όψιν καθώς αυτός είναι που σηκώνει το βάρος της χρήσης.

Στην παρούσα εργασία, προσπαθήσαμε μέσω της δημιουργίας ενός ιστοτόπου την επίλυση τέτοιων συστημάτων Captcha. Παρατηρήσαμε, ότι η επιτυχία επίλυσης είναι σχετικά μικρή αν αναλογιστούμε τα αποτελέσματα της προσπάθειας που συνοπτικά ήταν:

Είδος Capthca	**Περίπτωση Εφαρμογής**	**Επιτυχίες**	**Σύνολο Προσπαθειών**
SecureImage	Εγγραφή χρήστη	16/100	100
OpenCaptcha	Σύνδεση χρήστη	4/100	100
Snaphost	Επικοινωνία	21/100	100
reCaptcha	Σχολιασμός	0/100	100

Ο συντάκτης της εργασίας επιχείρησε μέσα από ένα συνδυασμό θεωρητικής και πρακτικής προσέγγισης να παρουσιάσει την τεχνολογία CAPTCHA στον αναγνώστη και να καταστήσει σαφές ότι μπορεί στην παρούσα φάση να αποτελεί μία αξιόπιστη λύση, αλλά θα πρέπει ταυτόχρονα να εξελίσσεται καθώς οι μηχανισμοί που προσπαθούν να την καταπολεμήσουν αυξάνονται και εξελίσσονται διαρκώς. Έγινε προσπάθεια να χρησιμοποιηθεί βιβλιογραφία από έγκριτα επιστημονικά περιοδικά των τελευταίων ετών, καθώς η τεχνολογία προχωρά και οι εξελίξεις τρέχουν. Ο

συντάκτης της εργασίας ευελπιστεί να ανταποκρίθηκε στις απαιτήσεις της εργασίας και στο επίπεδο της σχολής που ολοκληρώνει. Ταυτόχρονα ελπίζει να κάλυψε όλες τις πιθανές απορίες του αναγνώστη που κρατά το παρόν πόνημα στα χέρια του.

ΒΙΒΛΙΟΓΡΑΦΙΑ

- Ahn Nicholas J., Hopper Luis Von Blum, Manuel, and Langford. John (2000) The official CAPTCHA site

- Asirra - http://research.microsoft.com/en-us/um/redmond/projects/asirra/

- Chellapilla K. and Simard P.. (2004) Using machine learning to break visual human interaction proofs (HIPs). Advances in Neural Information Processing Systems, 17

- Chellapilla, K. Larson, K. Simard, P.Y. and Czerwinski M.. (2005) Building segmentation based human-friendly human interaction proofs (HIPs). Human Interactive Proofs, pages 1–26

- Chellapilla, K. Larson, K.. Simard , P and Czerwinski M.. (2005) Computers beat humans at single character recognition in reading based human interaction proofs (HIPs). In Proceedings of the Second Conference on Email and Anti-Spam, pages 21–22. Citeseer

- Chew M and Tygar JD. (2005) Image recognition CAPTCHAs. Information Security

- Chew Monica and. Tygar J. D, UC Berkeley (2004) Image Recognition CAPTCHAs In Proceedings of the 7th International Information Security Conference (ISC 2004), Springer, September 2004, pp. 268-279

- Coates, AL , Baird, HS and Fateman. RJ (2001) Pessimal print: a reverse Turing test. Proceedings of the Sixth International Conference on Document Analysis and Recognition

- E. Bursztein, S. Bethard, J. C. Mitchell, D. Jurafsky, and C. Fabry. How good are humans at solving CAPTCHAs? a large scale evaluation. In IEEE S&P ’10, 2010.

- Elson, Jeremy Douceur, John R. Howell, Jon and Saul. Jared (2007) Asirra: a CAPTCHA that exploits interest-aligned manual image categorization. In ACM Conference on Computer and Communications Security, pages 366–374. ACM, 2007.

- Frost. R (1998) Toward a strong phonological theory of visual word recognition: True issues and false trails. Psychological Bulletin

- Grainger J and Jacobs AM. (1996) Orthographic processing in visual word recognition: A multiple read-out model. Psychological review

- J. Yan and A. S. El Ahmad. Usability of CAPTCHAs or usability issues in CAPTCHA design. In SOUPS '08, pages 44–52, New York, NY, USA, 2008. ACM

- JoomlaWorks (2013, 05 03). Ανάκτηση από Simple Image Gallery: http://www.joomlaworks.gr/content/view/17/42/

- Lavrenko, V Rath, TM and Manmatha R (2004) . Holistic word recognition for handwritten historical documents. First International Workshop on Document Image Analysis for Libraries, 2004. Proceedings, 2004.

- Lillibridge, MD. Abadi, M. Bharat, K. and Border A.. (2001) Method for selectively restricting access to computer systems. US Patent 6,195,698

- Mariott, J., & Waring, E. (2010). pp.174, 183-184. The Official Joomla! Book. Addison Wesley.

- Milde B. (2010) On the security of re CAPTCHA Bachelor-Thesis from Darmstadt

- Mori and Malik J. (2003) Recognizing objects in adversarial clutter: Breaking a visual CAPTCHA

- Motoyama, Marti, Levchenko, Kirill, Kanich, Chris McCoy, Damon Voelker, Geoffrey M. and Savage, Stefan. (2010). Re: CAPTCHAs: understanding CAPTCHA-solving services in an economic context. In Proceedings of the 19th USENIX conference on Security (USENIX Security'10). USENIX Association, Berkeley, CA, USA, 28-28.

- Nagy, G Nartker, TA and SV Rice. (1999) Optical character recognition: An illustrated guide to the frontier. Proceedings of SPIE

- Nielsen Jakob. (2003) Usability 101: Introduction to Usability. Διαθέσιμο στο http://www.useit.com/alertbox/20030825.html.

- Rolko Juraj (2010) 3D CAPTCHA: CAPTCHA based on spatial perspective and human imagination Διαθέσιμο στο http://www.3dCAPTCHA.net/documents/3D_CAPTCHA.pdf

- Simard, P.Y. Szeliski, R.. Benaloh J,. Couvreur J, and Calinov I..(2003) Using character recognition and segmentation to tell computer from humans. Document Analysis and Recognition, 1:418

- Symantec. A captcha-solving service. http://www.symantec.com/connect/blogs/captcha-solving-service

- Wilkins. J (2009) Strong CAPTCHA guidelines v1. 2.

- Yan Jeff and Ahmad Ahmad Salah El. 2008. Usability of CAPTCHAs or usability issues in CAPTCHA design. In Proceedings of the 4th symposium on Usable privacy and security (SOUPS '08). ACM, New York, NY, USA, 44-52

- Content Management Systems - The History and the Future. (2013, 05 20). Ανάκτηση από Ezine Articles: http://ezinearticles.com/?Content-Management-Systems---The-History-and-the-Future&id=1665607.

- Content Management System: Ανάκτηση (2013, 05 20) από Wikipedia: http://en.wikipedia.org/wiki/Content_management_system.

- Joomla 1.5 API Reference. (2011,06 13). Ανάκτηση από Joomla: http://api.joomla.org/li_Joomla-Framework.html

- Model-view-controller. (2011,05 20). Ανάκτηση από Wikipedia: http://en.wikipedia.org/wiki/Model-view-controller.

- Porst, T. (2009). Joomla! 1.5 content administration . S.l.: Packet Pub Ltd.

- Severdia, R., & Crowder, K. (2010). Using Joomla . Sebastopol, Calif.: O'Reilly Media.

- GNU General Public Licence. (2013, 06 03). Ανάκτηση από Wikipedia: http://el.wikipedia.org/wiki/GNU_General_Public_License

- XAMPP. (2013, 05 03). Ανάκτηση από ApacheFriends: http://www.apachefriends.org/en/xampp.html

- P.J. Rousseeuw, A.M. Leroy, Robust Regression and Outlier Detection, Wiley-IEEE, 2003.

- K.S. Gursev Attacking CAPTCHAs for Fun and Profit, 2012
- T. Jennifer, J. Simsa, S.Hyde, L.V.Ahn Breaking Audio CAPTCHAs, 2008

ΔΙΚΤΥΑΚΕΣ ΠΗΓΕΣ

- http://www.phpcaptcha.org/documentation/quickstart-guide/ (SECURE IMAGE)
- http://webspamprotect.com/scripts_and_plugins.php (webspam protect)
- FreeCap - http://www.freecap.ru/eng/
- http://2glux.com/projects/jumi (jumi)
- Captcha - http://www.captcha.net/
- NuCAPTCHA - http://www.nuCAPTCHA.com/
- ReCaptcha - http://www.google.com/recaptcha/captcha
- Securimage - http://www.phpCAPTCHA.org/
- TextCAPTCHA - http://textCAPTCHA.com/
- Cryptographp - http://www.CAPTCHA.fr/
- W3 http://www.w3.org
- Webspamprotect - http://webspamprotect.com/