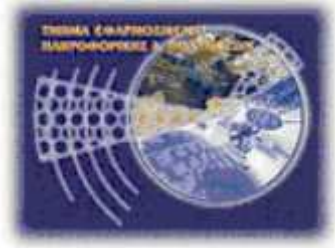




Τεχνολογικό Εκπαιδευτικό Ίδρυμα Κρήτης

Σχολή Τεχνολογικών Εφαρμογών

Τμήμα Εφαρμοσμένης Πληροφορικής και Πολυμέσων



Πτυχιακή εργασία

Τεχνικές CAPTCHA

Στέλλα Τσιγλάκη (Α.Μ. 2331)

Επιβλέπων Καθηγητής: Τριανταφυλλίδης Γεώργιος

Abstract

The spread of the Internet and the provision of free services from several websites had as a side effect their systematic abuse. To address the problem CAPTCHAs are widely used, in order to determine whether a request to a service is performed by a user or an automated program. The use of CAPTCHA is required in almost every website that allows users to create their own content or use its services. In this paper we examine the different types of CAPTCHA created to resist malicious solving attempts and analyze the process of using them in practice when creating a website.

Σύνοψη

Η εξάπλωση του διαδικτύου και η παροχή δωρεάν υπηρεσιών από αρκετές ιστοσελίδες είχε σαν παρενέργεια την συστηματική κατάχρηση τους. Για την αντιμετώπιση του προβλήματος χρησιμοποιούνται ευρέως τα CAPTCHA, με σκοπό να εξακριβώσουν αν μία αίτηση σε μία υπηρεσία γίνεται από έναν χρήστη ή από ένα αυτοματοποιημένο πρόγραμμα. Πλέον η χρήση CAPTCHA είναι απαραίτητη σχεδόν σε κάθε ιστοσελίδα που δίνει την δυνατότητα στον χρήστη να δημιουργήσει δικό του περιεχόμενο ή να χρησιμοποιήσει τις υπηρεσίες της. Σε αυτή την εργασία εξετάζονται τα διαφορετικά είδη CAPTCHA έχουν δημιουργηθεί ώστε να είναι ανθεκτικά στις κακόβουλες προσπάθειες λύσης και αναλύεται η διαδικασία χρήσης τους σε πρακτικό επίπεδο κατά την δημιουργία μίας ιστοσελίδας.

Περιεχόμενα

Abstract	3
Σύνοψη	4
1 Εισαγωγή	7
1.1 Γενικά.....	7
1.2 Σκοπός εργασίας	7
1.3 Μέθοδος ανάλυσης.....	7
1.4 Δομή εργασίας	8
2 Κακόβουλη χρήση υπηρεσιών	9
2.1 Γενικά.....	9
2.2 Κύριοι τομείς επιθέσεων.....	9
2.2.1 Δωρεάν υπηρεσίες ηλεκτρονικού ταχυδρομείου.....	9
2.2.2 Ηλεκτρονικές ψηφοφορίες	9
2.2.3 Ηλεκτρονικό marketing.....	10
2.2.4 Προώθηση ιστοσελίδων σε μηχανές αναζήτησης.....	10
2.2.5 Επίθεση λεξικού σε κωδικούς.....	10
2.3 Αρχικές προσπάθειες αντιμετώπισης του προβλήματος	11
2.3.1 Επιβολή παύσεων	11
2.3.2 Αναγνώριση μοναδικού υπολογιστή	12
3 Δοκιμασίες CAPTCHA	13
3.1 Εισαγωγή	13
3.2 Αιτίες χρήσης.....	13
3.3 Προβλήματα από την χρήση CAPTCHA.....	14
3.3.1 Άτομα με ειδικές ανάγκες.....	14
3.3.2 Προβλήματα ευχρηστίας	15
3.3.3 Μείωση conversion rate	15
4 Τεχνικές CAPTCHA	17
4.1 Εισαγωγή	17
4.2 Ερωτήσεις κειμένου	17
4.2.1 Μειονεκτήματα.....	19
4.2.2 Τυπικές επιθέσεις.....	19
4.3 CAPTCHA παραμόρφωσης γραμμάτων.....	20
4.3.1 Περιγραφή.....	20
4.3.2 Μειονεκτήματα.....	21

4.3.3	Τυπικές επιθέσεις.....	22
4.4	CAPTCHA Ήχου	23
4.4.1	Περιγραφή.....	23
4.4.2	Μειονεκτήματα	23
4.4.3	Τυπικές επιθέσεις.....	25
4.5	CAPTCHA Αναγνώρισης εικόνας.....	26
4.5.1	Περιγραφή.....	26
4.5.2	Μειονεκτήματα	28
4.5.3	Τυπικές επιθέσεις.....	29
4.6	Διαδραστικά CAPTCHA.....	29
4.6.1	Περιγραφή.....	29
4.6.2	Μειονεκτήματα	30
4.6.3	Τυπικές επιθέσεις.....	31
4.7	Σύνοψη	32
5	Παρουσίαση εφαρμογής.....	35
5.1	Γενικά.....	35
5.2	Τεχνολογίες υλοποίησης.....	35
5.2.1	HTML	35
5.2.2	JavaScript.....	36
5.2.3	PHP	37
5.3	Δομή εφαρμογής.....	38
5.3.1	Secureimage	38
5.3.2	Cryptographp.....	44
5.3.3	reCAPTCHA	48
5.3.4	Asirra	58
5.3.5	Sweetcaptcha	62
5.3.6	Civil Rights CAPTCHA	67
5.3.7	Text CAPTCHA.....	70
6	Συμπεράσματα	74
7	Βιβλιογραφία.....	76

1 Εισαγωγή

1.1 Γενικά

Τα CAPTCHA είναι ένα είδος αντίστροφου τεστ Turing με στόχο να επιβεβαιώσει ότι η απάντηση σε αυτό προέρχεται από έναν άνθρωπο και όχι από έναν υπολογιστή. Στην πιο διαδεδομένη τους μορφή είναι παραμορφωμένες εικόνες κειμένου και η συμπλήρωσή τους είναι απαραίτητη σε πάρα πολλές υπηρεσίες, από την υποβολή ενός σχολίου για κάποιο κείμενο και την δημιουργία ενός λογαριασμού ηλεκτρονικού ταχυδρομείου μέχρι την προσπάθεια ανάκτησης κωδικού σε υπηρεσίες ηλεκτρονικού εμπορίου, με σκοπό να σταματήσουν τις αυτοματοποιημένες δέσμες ενεργειών που χρησιμοποιούνται για την κακόβουλη χρήση αυτών των υπηρεσιών.

Το κύριο χαρακτηριστικό τους είναι ότι παρουσιάζουν ένα πρόβλημα η επίλυση του οποίου είναι εύκολη για τους ανθρώπους αλλά αρκετά δύσκολη έως αδύνατη για τους υπολογιστές. Αυτόματα με την εισαγωγή τους ξεκίνησε ένας αγώνας δρόμου με σκοπό την ανάπτυξη τεχνολογίας για την επίλυσή τους και την χρήση πιο πολύπλοκων τεστ για την αχρήστευση αυτών των τεχνολογιών. Αυτό έχει σαν αποτέλεσμα, πέρα από την εξέλιξη των βασικών οπτικών δοκιμασιών, την ανάπτυξη νέων ειδών CAPTCHA που βασίζονται σε διαφορετικές τεχνικές.

Η δημιουργία ενός συστήματος CAPTCHA είναι αρκετά δύσκολη και πολύπλοκη από άποψη γνώσεων, επένδυσης σε χρόνο ανάπτυξης και αρκετές φορές σε υπολογιστική ισχύ για την δημιουργία του τεστ. Ο συνηθέστερος δρόμος είναι η χρήση μίας έτοιμης λύσης αλλά πλέον ακόμη και αυτό χρειάζεται ιδιαίτερη προσπάθεια.

Η πληθώρα λύσεων που υπάρχουν μαζί με την ανάγκη χρήσης CAPTCHA στην πλειοψηφία των σελίδων δημιουργεί ένα αρκετά δύσκολο τοπίο για τους δημιουργούς των ηλεκτρονικών υπηρεσιών, καθώς χρειάζεται να επιλέξουν μία αποτελεσματική λύση και έπειτα να την ενσωματώσουν στον κώδικά τους. Αυτό οφείλεται στις διαφορετικές τεχνολογίες που χρησιμοποιεί κάθε τεστ τόσο για την υλοποίησή του όσο και για την χρήση του.

1.2 Σκοπός εργασίας

Αν και η χρήση CAPTCHA είναι πλέον αναγκαία, η δυσκολία επιλογής ανάμεσα στις διαθέσιμες προτάσεις και χρήσης μίας από αυτές είναι αρκετά μεγάλη. Με αυτό το σκεπτικό αφού αναλύσουμε κάποιες βασικές έννοιες σχετικά με αυτά εξετάζουμε διαφορετικές προσεγγίσεις του προβλήματος και σχολιάζουμε τα πλεονεκτήματα και τα μειονεκτήματά τους.

Επιπλέον σε πρακτικό επίπεδο αναλύουμε πως οι έτοιμες λύσεις μπορούν να ενσωματωθούν σε μία ιστοσελίδα παραθέτοντας τις λεπτομέρειες και τις δυσκολίες υλοποίησης για κάθε μία από αυτές.

1.3 Μέθοδος ανάλυσης

Με στόχο να ικανοποιηθεί ο προηγούμενος στόχος κατασκευάστηκε μία ιστοσελίδα η οποία περιέχει διαφορετικές φόρμες HTML, κάθε μία από τις οποίες χρησιμοποιεί τις

διαφορετικές λύσεις. Με αυτό τον τρόπο είναι εφικτή η παράθεση των λεπτομερειών υλοποίησης για κάθε μία λύση και η σύγκριση των διαφορετικών λύσεων μεταξύ τους.

1.4 Δομή εργασίας

Η παρούσα εργασία ξεκινάει παραθέτοντας στο δεύτερο κεφάλαιο το πρόβλημα που καλούνται να αναλύσουν τα CARTCHA και τις εφαρμογές που έχουν.

Στο τρίτο κεφάλαιο εισάγεται η έννοια των CARTCHA παραθέτοντας τα γενικά πλεονεκτήματα και μειονεκτήματά τους.

Στο τέταρτο κεφάλαιο αναφέρουμε τα διαφορετικά είδη CARTCHA και τον τρόπο προσέγγισης του προβλήματος από τις διαφορετικές υλοποιήσεις, ενώ στο πέμπτο κεφάλαιο παραθέτουμε τις τεχνικές λεπτομέρειες για κάθε μία από αυτές.

Στο πέμπτο κεφάλαιο αναλύεται η εφαρμογή που αναπτύχθηκε και οι διαφορετικές υπηρεσίες που χρησιμοποιήθηκαν.

Τέλος στο έκτο κεφάλαιο παραθέτουμε τα συμπεράσματα στα οποία καταλήξαμε μετά την θεωρητική και την πρακτική μελέτη των διαφορετικών τεχνολογιών.

2 Κακόβουλη χρήση υπηρεσιών

2.1 Γενικά

Η αύξηση της χρήσης του διαδικτύου που ξεκίνησε την αρχή της δεκαετίας του 1990 και από τη στιγμή που ξεκίνησε να προσελκύει ένα αρκετά μεγάλο πλήθος χρηστών εμφανίστηκαν και τα πρώτα κρούσματα εκμετάλλευσής του από επιτηδείους με σκοπό την αποκόμιση χρηματικών οφελών. Κοινό χαρακτηριστικό τέτοιων προσπαθειών είναι η χρήση αυτοματοποιημένων εργαλείων για την μαζική πραγματοποίηση ενεργειών που σε έναν κανονικό χρήστη θα έπαιρναν αρκετό χρόνο.

2.2 Κύριοι τομείς επιθέσεων

Παρακάτω παρατίθενται οι κυριότεροι τομείς που παρουσιάστηκαν τέτοια προβλήματα.

2.2.1 Δωρεάν υπηρεσίες ηλεκτρονικού ταχυδρομείου

Η πρώτη μορφή κατάχρησης ήταν η μαζική αποστολή διαφημιστικών ηλεκτρονικών μηνυμάτων χωρίς την συγκατάθεση των παραληπτών (spam) που γρήγορα πήρε μεγάλες διαστάσεις και άρχισε να αντιμετωπίζεται εχθρικά από τις εταιρείες του διαδικτύου με σκοπό της εξάλειψής της. Αυτό κατέστησε την αποστολή μηνυμάτων spam οικονομικά ασύμφορη, καθώς ο αποστολέας είναι απαραίτητο να διατηρεί την δική του υποδομή αποστολής email.

Όμως η πρόοδος της τεχνολογίας άρχισε να κάνει εφικτή την παροχή δωρεάν υπηρεσιών ηλεκτρονικού ταχυδρομείου στους χρήστες. Οι μέχρι τότε spammers εκμεταλλεύτηκαν τις υπηρεσίες αυτές καθώς είχαν το πλεονέκτημα ότι δεν χρειαζόντουσαν συντήρηση των υποδομών (αφού προσφέρονταν δωρεάν) και δεν ήταν εύκολο να μπλοκαριστούν γιατί φιλοξενούνται σε εξυπηρετητές που θεωρούνται ασφαλείς.

Βέβαια ένας μικρός αριθμός λογαριασμών σε μία τέτοια υπηρεσία δεν αρκεί για την αποστολή του όγκου του μηνυμάτων που χρειάζονται, καθώς οι μεμονωμένοι λογαριασμοί γίνονται γρήγορα αντιληπτοί και μπλοκάρονται. Αυτό κατέστησε απαραίτητη την μαζική δημιουργία λογαριασμών ώστε να αναπληρώνονται αυτοί που κλείνουν με την χρήση δεσμών ενεργειών.

2.2.2 Ηλεκτρονικές ψηφοφορίες

Πολλές σελίδες περιέχουν μαζί με το περιεχόμενό τους και ηλεκτρονικές ψηφοφορίες. Πολλές φορές αυτές περιέχουν ερωτήσεις που διχάζουν το κοινό και μπορεί ένα άτομο με μία συγκεκριμένη άποψη να προσπαθήσει να αλλοιώσει το αποτέλεσμα της ψηφοφορίας ώστε να υποστηρίξει την θέση του.

Αν η ιστοσελίδα που διοργανώνει την ψηφοφορία έχει πολλούς αναγνώστες και κατά συνέπεια η ψηφοφορία έχει αρκετά μεγάλο δείγμα οι προσπάθειες ενός ατόμου δεν μπορούν να αλλοιώσουν σημαντικά το αποτέλεσμα. Όμως η χρήση ηλεκτρονικών εργαλείων που μπορούν να παράγουν χιλιάδες ψήφους στον χρόνο που θα έπαιρνε σε έναν άνθρωπο να ψηφίσει μία φορά μπορεί να αλλάξει τελείως το αποτέλεσμα της πληροφορίας

αχρηστεύοντας το περιεχόμενό της. Ο μοναδικός τρόπος για την διασφάλιση του αποτελέσματος είναι η εξασφάλιση ότι ψηφίζουν μόνο άνθρωποι.

2.2.3 Ηλεκτρονικό marketing

Κάποιες εταιρείες στην προσπάθειά τους να προωθήσουν τα προϊόντα τους χρησιμοποιούν αθέμιτες τακτικές. Μία από αυτές είναι η ανάρτηση σε forum και blog περιεχομένου που να τα εξυμνεί. Όμως η δημιουργία τέτοιων τόπων αφιερωμένων σε αυτό δεν προσφέρει τίποτα γιατί κανείς δεν ενδιαφέρεται να επισκεφτεί μία σελίδα που ξέρει ότι περιέχει μόνο διαφημιστικό υλικό. Ο στόχος είναι η ανάρτηση τέτοιου υλικού σε άλλες ευυπόληπτες σελίδες ώστε να δίνουν την εντύπωση ότι πρόκειται για ειλικρινή σχόλια χρηστών. Ιδιαίτερα συνηθισμένος στόχος είναι τα σχόλια που μπορούν να δημοσιεύσουν οι χρήστες σχετικά με άρθρα, καθώς στις περισσότερες περιπτώσεις δεν χρειάζεται η δημιουργία λογαριασμού ή η επικύρωση των στοιχείων του χρήστη.

Η δημιουργία τέτοιων σχολίων από απλούς χρήστες είναι περιορισμένης χρησιμότητας γιατί μπορούν να δημοσιεύσουν πεπερασμένο αριθμό σχολίων και δεν μπορούν να επηρεάσουν μεγάλο μέρος αναγνωστών, δεδομένου ότι οι περισσότεροι από αυτούς θα καταλάβουν ότι το μήνυμα είναι πλαστό και θα το αγνοήσουν. Αντίθετα η χρήση αυτοματοποιημένων εργαλείων προσφέρει πολύ μεγαλύτερη ακτίνα δράσης.

2.2.4 Προώθηση ιστοσελίδων σε μηχανές αναζήτησης

Εξαιτίας την άναρχης αρχιτεκτονικής του διαδικτύου οι περισσότεροι χρήστες γνωρίζουν ελάχιστες διευθύνσεις ιστοσελίδων, κυρίως αυτών που έχουν συγκεκριμένο περιεχόμενο που ανταποκρίνεται στα ενδιαφέροντά τους. Αν όμως ενδιαφέρονται να βρουν πληροφορίες σχετικά με κάτι νέο δεν έχουν άλλη επιλογή παρά να καταφύγουν στις μηχανές αναζήτησης.

Μία αναζήτηση μίας συγκεκριμένης φράσης σε μία μηχανή αναζήτησης παράγει συνήθως εκατομμύρια αποτελέσματα. Η ποιότητα των αποτελεσμάτων δεν εξαρτάται από το πλήθος τους, καθώς σπάνια ένας χρήστης εξετάζει παραπάνω από τα πρώτα δέκα, αλλά από τον τρόπο με τον οποίο τα ταξινομούν. Οι μηχανές αναζήτησης χρησιμοποιούν συγκεκριμένους αλγορίθμους ώστε να βαθμολογήσουν πόσο σχετίζεται μία συγκεκριμένη σελίδα με τον όρο που έψαξε ο χρήστης. Αυτή η βαθμολογία εξαρτάται από πολλούς παράγοντες, ένας από τους οποίους είναι το πόσο συχνά αναφέρεται (μέσω υπερσυνδέσεων) σε άλλες σελίδες.

Για την άνοδο μίας σελίδας στη λίστα των αποτελεσμάτων βοηθάει λοιπόν η αναφορά της σε άλλες σελίδες. Ένας τρόπος για να γίνει αυτό είναι, όπως είδαμε στην προηγούμενη παράγραφο, η ανάρτηση σχολίων σε άλλες σελίδες με αυτοματοποιημένο τρόπο, οι οποίες να περιέχουν συνδέσμους στις σελίδες που θέλουν να προωθήσουν.

2.2.5 Επίθεση λεξικού σε κωδικούς

Κάποιες φορές στόχος μίας επίθεσης είναι η απόκτηση πρόσβασης στον ηλεκτρονικό λογαριασμό κάποιων χρηστών. Η σοβαρότητα των επιπτώσεων μίας τέτοιας επίθεσης μπορεί να διαφέρει, καθώς εκτινεται από την πρόσβαση σε έναν λογαριασμό ηλεκτρονικού ταχυδρομείου με σκοπό την αποστολή spam emails, που προκαλεί ενόχληση στον χρήστη και τους παραλήπτες των μηνυμάτων, μέχρι και την πρόσβαση σε τραπεζικούς

λογαριασμούς ή ιατρικές σελίδες που μπορούν να προκαλέσουν μεγάλη οικονομική ζημιά και κλοπή προσωπικών δεδομένων.

Ο πιο διαδεδομένος τρόπος εξακρίβωσης της ταυτότητας ενός χρήστη και της παροχής πρόσβασης στον λογαριασμό που έχει σε κάποια διαδικτυακή υπηρεσία είναι μέσω της εισαγωγής ενός μυστικού κωδικού. Όμως αρκετοί χρήστες επιλέγουν αδύναμους κωδικούς που μπορεί να είναι μία παραλλαγή του ονόματός τους, τα γενέθλια τους ή κάποια λέξη.

Ένας τρόπος να αποκτήσει κάποιος πρόσβαση στον λογαριασμό ενός χρήστη είναι να προσπαθήσει να μαντέψει τον κωδικό του διαλέγοντας λέξεις από ένα λεξικό, το οποίο πολλές φορές περιέχει λέξεις επιλεγμένες εξαιτίας της συχνότητας που χρησιμοποιούνται σε κωδικούς. Όμως συνήθως για την εύρεση ενός κωδικού απαιτούνται χιλιάδες προσπάθειες, οι οποίες είναι απαραίτητο να γίνουν από αυτόματα εργαλεία.

2.3 Αρχικές προσπάθειες αντιμετώπισης του προβλήματος

Το κοινό χαρακτηριστικό που έχουν όλες οι παραπάνω κακόβουλες χρήσεις που αναφέρθηκαν είναι η χρήση αυτοματοποιημένων εργαλείων που εκτελούν μία χρονοβόρα για έναν άνθρωπο διαδικασία αρκετές φορές, από εκατοντάδες έως και χιλιάδες. Όταν παρουσιάστηκαν τα πρώτα κρούσματα οι διαχειριστές των ιστοσελίδων προσπάθησαν να τα αντιμετωπίσουν χρησιμοποιώντας διάφορες τακτικές.

2.3.1 Επιβολή παύσεων

Η πρώτη μέθοδος αντιμετώπισης ήταν η επιβολή συγκεκριμένων χρονικών περιθωρίων μέσα στα οποία μπορούσε να γίνει κάποια ενέργεια. Το σκεπτικό είναι ότι αν και ένας υπολογιστής μπορεί να εκτελέσει την διαδικασία πολλές φορές κάθε δευτερόλεπτο, ένας άνθρωπος μπορεί να το κάνει μόνο μία φορά κάθε αρκετά δευτερόλεπτα. Λαμβάνοντας λοιπόν υπόψη την διαφορά στον χρόνο αντίδρασης επιβλήθηκαν περιορισμοί που δεν επιτρέπουν την επανάληψή της άμεσα.

Για παράδειγμα ένας χρήστης που δοκιμάζει να εισάγει τον κωδικό του και δίνει λάθος κωδικό δεν επιτρέπεται να δοκιμάσει άμεσα, αλλά πρέπει να περιμένει μερικά δευτερόλεπτα, ενώ από τη στιγμή που κάποιος κάνει ένα σε ένα blog δεν θα θέλει να δημοσιεύσει και δεύτερο για μερικά. Αυτή η τακτική εμποδίζει την επανάληψη της διαδικασίας αρκετά συχνά ώστε να είναι χρήσιμη στον επιτιθέμενο.

Δυστυχώς αυτή η τεχνική έχει δύο μειονεκτήματα. Το πρώτο είναι πως μπορεί να είναι ενοχλητική για τους έγκυρους χρήστες, καθώς υπάρχουν σενάρια χρήσης που θέλουν να κάνουν μία ενέργεια αρκετά γρήγορα. Κάποιος που έβαλε λάθος τον κωδικό του θέλει να ξαναδοκιμάσει αμέσως για να μπει στον λογαριασμό του, ενώ ένας χρήστης που έκανε ένα σχόλιο σε ένα άρθρο μπορεί να θέλει να αναδιατυπώσει άμεσα την άποψή του για να αποφύγει κάποια παρεξήγηση.

Επιπλέον για τους δημιουργούς των κακόβουλων επιθέσεων η συγκεκριμένη τεχνική δεν δημιουργεί ιδιαίτερα προβλήματα γιατί αρκεί να αλλάξουν τις παραμέτρους των συστημάτων τους. Για παράδειγμα μπορεί κάποιος να επιχειρεί να μαντέψει τον κωδικό ενός χρήστη κάνοντας 10 προσπάθειες κάθε δευτερόλεπτο. Αν η υπηρεσία που

προσφέρει το ηλεκτρονικό ταχυδρομείο βάλει ένα όριο 10 δευτερολέπτων ανά προσπάθεια, τότε ο επιτιθέμενος μπορεί να τρέξει 10 τέτοια προγράμματα για διαφορετικούς χρήστες, ρυθμισμένα να δοκιμάζουν έναν νέο κωδικό για τον καθένα κάθε 10 δευτερόλεπτα. Καθώς και στις δύο περιπτώσεις κάνει 100 προσπάθειες κάθε 10 δευτερόλεπτα δεν έχει κάποια ουσιαστική απώλεια.

2.3.2 Αναγνώριση μοναδικού υπολογιστή

Η δεύτερη τακτική αποτροπής βασίζεται στη λογική ότι ένας υπολογιστής μπορεί να κάνει μία ενέργεια μόνο μία φορά (ή μόνο μία φορά σε ένα συγκεκριμένο χρονικό διάστημα, συνδυάζοντας το με την τακτική της προηγούμενης ενότητας). Για παράδειγμα σε μία ηλεκτρονική ψηφοφορία ένας υπολογιστής μπορεί να ψηφίσει μόνο μία φορά, ενώ από έναν υπολογιστή δεν είναι αρκετά πιθανό να ανοιχθεί πάνω από ένας λογαριασμός ηλεκτρονικού ταχυδρομείου την ημέρα.

Αν και η ιδέα είναι λογική δεν είναι εφικτό να εκτελεστεί. Για την εξακρίβωση της ταυτότητας ενός υπολογιστή μπορούν να χρησιμοποιηθούν μόνο δύο τρόποι, η χρήση cookies ή ενός session key και η παρακολούθηση της διεύθυνσης IP.

Η χρήση ενός μοναδικού session key και η συσχέτισή του με κάποιον υπολογιστή δεν είναι εφικτή γιατί αυτό θα πρέπει να αποθηκεύεται στον επισκέπτη. Ένας νόμιμος χρήστης δεν πρόκειται να το πειράξει αλλά δεν πρόκειται και να έχει παραβατική συμπεριφορά. Αντίθετα ένας κακόβουλος χρήσης μπορεί πολύ εύκολα να το σβήσει ή τα το τροποποιήσει ώστε να εξυπηρετήσει τον σκοπό του.

Η χρήση της διεύθυνσης IP είναι ακόμα πιο δύσκολη και μη πρακτική. Αν και θεωρητικά κάθε υπολογιστής που συνδέεται στο διαδίκτυο έχει μία μοναδική διεύθυνση πρακτικά αυτό έχει πάψει να ισχύει εξαιτίας της προόδου της τεχνολογίας. Σε πολλές περιπτώσεις οι χρήστες δεν έχουν την στατική τους διεύθυνση αλλά αυτή αλλάζει κατά διαστήματα από τον πάροχο διαδικτυακής υπηρεσίας (ISP) που χρησιμοποιούν. Ακόμα χειρότερα μπορεί μία διεύθυνση να αντιστοιχεί σε έναν δρομολογητή μέσω του οποίου να συνδέονται πολλοί χρήστες ενός δικτύου. Στα τοπικά δίκτυα μεγάλων οργανισμών χιλιάδες υπολογιστές έχουν πρόσβαση στο διαδίκτυο μέσω μίας μόνο διεύθυνσης, κάτι που δεν μπορεί να εξακριβωθεί.

Τα παραπάνω δεν σταματούν κάποιους διαχειριστές ιστοσελίδων από το να απαγορεύουν την πρόσβαση σε συγκεκριμένες διευθύνσεις IP ή και σε ολόκληρα IP ranges στις οποίες παρατηρείται επανειλημμένα παραβατική συμπεριφορά, περιορίζοντας με αυτό τον τρόπο την πρόσβαση πάρα πολλών χρηστών στο διαδίκτυο. Και καθώς τα τελευταία χρόνια με την μεγάλη εξάπλωση του malware και των δικτύων zombie υπολογιστών υπάρχει περίπτωση να γίνονται επιθέσεις από τελείως άσχετες μεταξύ τους διευθύνσεις τέτοιες τακτικές δεν έχουν σε καμία περίπτωση το επιθυμητό αποτέλεσμα.

3 Δοκιμασίες CAPTCHA

3.1 Εισαγωγή

Το ακρωνύμιο CAPTCHA βασίζεται στην λέξη capture (σύλληψη) και σημαίνει **C**ompletely **A**utomated **P**ublic **T**uring test to tell **C**omputers and **H**umans **A**part (Δημόσιο Πλήρως Αυτοματοποιημένο τεστ Turing για τον διαχωρισμό ανθρώπων και υπολογιστών). Τα CAPTCHA δημιουργήθηκαν για να εγγυηθούν ότι αυτός που απαντάει είναι ένας άνθρωπος. Μπορεί να χαρακτηριστεί σαν αντίστροφο τεστ Turing, γιατί ενώ το αυθεντικό τεστ το κάνει ένας άνθρωπος σε μία μηχανή, το τεστ CAPTCHA το κάνει μία μηχανή σε έναν άνθρωπο.

Όλοι οι τρόποι κατάχρησης των υπηρεσιών που αναλύσαμε παραπάνω χρειάζονται την χρήση αυτοματοποιημένων εργαλείων για να υλοποιηθούν, τα οποία κάνουν συνεχώς διαδικασίες που είναι εξαιρετικά χρονοβόρες για ένα άτομο ή μία ομάδα ατόμων. Στόχος του CAPTCHA είναι να δημιουργήσει μία ερώτηση που η λύση της είναι αρκετά απλή για έναν άνθρωπο αλλά βασίζεται σε ένα πρόβλημα που δεν λύνεται ή λύνεται εξαιρετικά δύσκολα από έναν υπολογιστή.

3.2 Αιτίες χρήσης

Το μέγεθος του προβλήματος του spam έχει επεκταθεί σε μεγάλο βαθμό, ειδικά με την διάδοση των sites που δίνουν βάρος στην παρουσίαση περιεχομένου που δημιουργούν οι χρήστες. Το akismet, που είναι ένα σύστημα αντιμετώπισης των αυτοματοποιημένων μηνυμάτων για το σύστημα διαχείρισης περιεχομένου wordpress σταματάει καθημερινά περισσότερα από 18 εκατομμύρια σχόλια από ψεύτικους χρήστες, ενώ διάφορες υπηρεσίες παρακολούθησης περιεχομένου υπολογίζουν ότι το ποσοστό των μηνυμάτων spam ξεπερνάει το 90% του συνόλου του περιεχομένου που δημιουργείται καθημερινά.

Για τον διαχειριστή μίας ιστοσελίδας τα προβλήματα που δημιουργούνται είναι αρκετά σοβαρά:

- Το πραγματικό περιεχόμενο χάνεται μέσα στην πληθώρα του spam.
- Οι χρήστες έχουν αρκετά δύσκολη πρόσβαση στο περιεχόμενο των άλλων.
- Έχοντας την πλειοψηφία των λογαριασμών της σελίδας να ανήκει σε ψεύτικους χρήστες αχρηστεύονται οι προσπάθειες επικοινωνίας με τους πραγματικούς χρήστες και η προώθηση διαφημίσεων ή άλλου υλικού σε αυτούς.
- Είναι πιθανό να χρησιμοποιηθούν ψεύτικοι λογαριασμοί για παράνομους σκοπούς, δημιουργώντας νομικά προβλήματα στη σελίδα.
- Με το πέρασμα του χρόνου οι χρήστες θα αρχίσουν να εγκαταλείπουν τέτοιες σελίδες, καθώς η εμπειρία χρήσης και η ποιότητα του περιεχομένου θα έχουν πέσει σε απαράδεκτα επίπεδα.

Όμως η παρακολούθηση όλων των νέων λογαριασμών που δημιουργούνται και του περιεχομένου που αυτοί δημιουργούν δεν είναι πρακτική. Οι διαχειριστές μικρών σελίδων τυπικά αφιερώνουν ένα μικρό μέρος του χρόνου τους σε αυτές και δεν έχουν την δυνατότητα να ασχολούνται με οποιαδήποτε ενέργεια κάθε χρήστη για να σταματάνε τα αυτοματοποιημένα εργαλεία.

Το πρόβλημα γίνεται ακόμη μεγαλύτερο όσο περισσότερο πληθαίνουν οι χρήστες μίας σελίδας, καθώς το περιεχόμενο που αυτοί δημοσιεύουν αυξάνεται εκθετικά. Σε διάφορες έρευνες που έχουν γίνει έχει υπολογιστεί ότι για τις μεγαλύτερες σελίδες του internet όπως το facebook και το youtube, το κόστος παρακολούθησης του περιεχομένου που αναρτούν οι χρήστες ανέρχεται σε δισεκατομμύρια δολάρια, πολύ μεγαλύτερο από τον κύκλο εργασιών τους.

Τα παραπάνω έχουν καταστήσει τα CAPTCHA την αυτονόητη και πολλές φορές αυτόματη επιλογή στα εργαλεία αντιμετώπισης του spam. Σχεδόν οποιαδήποτε φόρμα καλείται να συμπληρώσει ένας χρήστης είναι συνηθισμένο να έχει και ένα τέτοιου είδους τεστ, ενώ το reCAPTCHA, μία από τις πιο δημοφιλείς υπηρεσίες CAPTCHA, υπολογίζει ότι καθημερινά λύνονται 200 εκατομμύρια CAPTCHA. Όμως η χρήση τους σε τέτοια κλίμακα γίνεται χωρίς να ληφθούν υπόψη κάποια σοβαρά μειονεκτήματά τους.

3.3 Προβλήματα από την χρήση CAPTCHA

Τα CAPTCHA έχουν φτιαχτεί με σκοπό την άμεση αντιμετώπιση των spammers, ενώ κατά τον σχεδιασμό και την χρήση τους οι πραγματικοί χρήστες αγνοούνται. Αυτό γίνεται γιατί κυριαρχεί η υπόθεση ότι οι πραγματικοί χρήστες δεν επηρεάζονται από αυτά, μία υπόθεση που είναι λαθεμένη σε μεγάλο βαθμό.

Αρχικά πρέπει να σημειωθεί ότι τα πλεονεκτήματα που πολλοί θεωρούν ότι έχει η χρήση των CAPTCHA, δηλαδή να σταματάει τα αυτοματοποιημένα εργαλεία που έχουν στην διάθεσή τους οι spammers, δεν υφίστανται, ή έστω ισχύουν σε πάρα πολύ μικρό βαθμό. Έρευνες έχουν δείξει ότι τα περισσότερα από αυτά, σε ποσοστά από 85% έως και 99% μπορούν να λυθούν χρησιμοποιώντας σχετικά απλές τεχνικές OCR. Πρόκειται για ένα πρόβλημα που δεν μπορεί να αντιμετωπιστεί με την βελτίωση της τεχνολογίας.

Ο σκοπός του CAPTCHA είναι να παρουσιάσει ένα πρόβλημα που είναι απλό στη λύση του για έναν άνθρωπο και δύσκολο να λυθεί από μία μηχανή. Με το πέρασμα του χρόνου όμως οι μηχανές αποκτούν όλο και περισσότερες δυνατότητες, και το τι είναι δύσκολο να λυθεί από αυτές δεν είναι σαφώς ορισμένο. Για να αντιμετωπιστεί αυτό παρουσιάζονται προβλήματα που είναι ολοένα και πιο πολύπλοκα, όμως με αυτό τον τρόπο γίνεται δύσκολη η λύση τους από τους απλούς ανθρώπους.

3.3.1 Άτομα με ειδικές ανάγκες

Εξ' ορισμού ο σκοπός του CAPTCHA είναι να μην είναι προφανής ο τρόπος χρήσης μίας σελίδας, συνήθως την συμπλήρωσης μίας φόρμας. Αυτό, πέρα από την προφανή μείωση στην ευχρηστία της σελίδας στην οποία χρησιμοποιείται, δημιουργεί μεγάλα προβλήματα σε χρήστες με ειδικές ανάγκες, και ειδικά σε άτομα που μπορεί να έχουν μειωμένη όραση ή τύφλωση, σε ανθρώπους με προβλήματα βαρηκοΐας και κώφωσης και ανάλογα με το είδος και σε άτομα με κινητικά προβλήματα.

Ο συνηθισμένος τρόπος αντιμετώπισης αυτού του φαινομένου είναι η προσφορά και μίας εναλλακτικής μορφής CAPTCHA, πχ αυτά που βασίζονται στην άριστη όραση του χρήστη προσφέρουν και έναν τύπο ακουστικής ερώτησης, αλλά πρόκειται για ένα ημίμετρο που απέχει πάρα πολύ από το να είναι ικανοποιητικό. Βασίζεται πάλι στην λανθασμένη υπόθεση πως ο χρήστης δεν θα έχει πολλαπλές αναπηρίες και επιπλέον κάποια είδη απλά

δεν είναι πρακτικά στην χρήση τους, όπως για παράδειγμα τα ακουστικά CAPTCHA σε μέρη όπως βιβλιοθήκες και αναγνωστήρια όπου συνήθως επιβάλλεται απόλυτη ησυχία ή σε θορυβώδη μέρη.

3.3.2 Προβλήματα ευχρηστίας

Αλλά ακόμη και αν αγνοήσουμε τα προβλήματα των χρηστών με ειδικές ανάγκες τα CAPTCHA συνεχίζουν να βασίζονται σε μία λαθεμένη λογική γιατί υποχρεώνουν τους χρήστες να τα λύσουν για να συνεχίσουν. Το spam δεν είναι πρόβλημα του χρήστη αλλά της εταιρείας που παρέχει την υπηρεσία. Το να απαιτείται από τον χρήστη να λύσει αυτό το πρόβλημα είναι αλαζονικό. Ένα μέσο CAPTCHA χρειάζεται περίπου δέκα δευτερόλεπτα για να λυθεί, ενώ πλέον έχουν γίνει τόσο δύσκολα που για κάποια από αυτά χρειάζονται δύο ή και τρεις προσπάθειες για να λυθούν.

Σαν λύση σε αυτό προτάθηκε η παραγωγική χρήση του χρόνου του χρήστη, με σκοπό να μειώσει την ενόχλησή του. Το σημαντικότερο παράδειγμα είναι η υπηρεσία reCAPTCHA, η οποία παρουσιάζει στους χρήστες δύο λέξεις από βιβλία που βρίσκονται στην διαδικασία ψηφιοποίησης τους. Καμία από τις δύο δεν έχει περάσει με επιτυχία από το πρόγραμμα OCR που αρχικά χρησιμοποιήθηκε για την αναγνώρισή τους και έτσι πέρασαν σαν CAPTCHA στους χρήστες ώστε να λυθούν από ανθρώπους.

Το σκεπτικό πίσω από αυτή την υπηρεσία είναι πως οι χρήστες δεν θα έχουν πρόβλημα να αφιερώσουν χρόνο και κόπο για να λύσουν το τεστ, καθώς με αυτόν τον τρόπο προσφέρουν στο σύνολο. Όμως αγνοεί το γεγονός ότι ο χρήστης δεν επέλεξε ο ίδιος να προσφέρει σε αυτή την προσπάθεια, αλλά του επιβλήθηκε. Ακόμη χειρότερα, οι λέξεις τις οποίες καλείται να αναγνωρίσει επιλέχθηκαν με κριτήριο το ότι δεν μπορούν να αναγνωριστούν από τα κορυφαία προγράμματα OCR, δηλαδή επειδή είναι αρκετά δυσανάγνωστες και απαιτούν μεγάλη προσπάθεια για να λυθούν.

Ένας χρήστης ο οποίος επιθυμεί να χρησιμοποιήσει μία υπηρεσία δεν μπορεί να αντιμετωπίζεται με αυτόν τον τρόπο, καθώς αυτή η δυσκολία δημιουργεί πάρα πολύ αρνητικές εντυπώσεις και σε ορισμένες επιπτώσεις μπορεί να διώξει τελείως τον χρήστη από την σελίδα. Πρόκειται για ένα φαινόμενο που έχει μελετηθεί και πλέον λαμβάνεται υπόψη κατά το SEO μίας σελίδας.

3.3.3 Μείωση conversion rate

Ένας από τους πιο σημαντικούς δείκτες που χαρακτηρίζουν το πόσο πετυχημένη είναι μία ιστοσελίδα είναι το ποσοστό μετατροπής (conversion rate). Πρόκειται για το ποσοστό των χρηστών που ολοκληρώνουν μία συγκεκριμένη ενέργεια σε σχέση με τον αριθμό των χρηστών που ξεκινάνε να κάνουν αυτή την ενέργεια. Για παράδειγμα μπορεί να πρόκειται για το πόσοι χρήστες ολοκληρώνουν μία ηλεκτρονική παραγγελία ή πόσοι χρήστες εγγράφονται σε μία ιστοσελίδα.

Πρόκειται για έναν πολύ σημαντικό δείκτη, οι διακυμάνσεις του οποίου μεταφράζονται άμεσα σε διακυμάνσεις σε κέρδη για τις εταιρείες, είτε μέσω διαφημίσεων είναι μέσω πωλήσεων μέσω του site. Αυτός ο δείκτης πλήττεται άμεσα από την χρήση captcha, η οποία υπολογίζεται πως μειώνει το ποσοστό μετατροπής κατά 3 με 4%. Δηλαδή η υπηρεσία είχε πείσει τους χρήστες να την χρησιμοποιήσουν και έπειτα έδιωξε κάποιους

από αυτούς με τον τρόπο με τον οποίο επιχείρησε να τους μεταβιβάσει την λύση στο πρόβλημα spam που αντιμετωπίζει.

4 Τεχνικές CAPTCHA

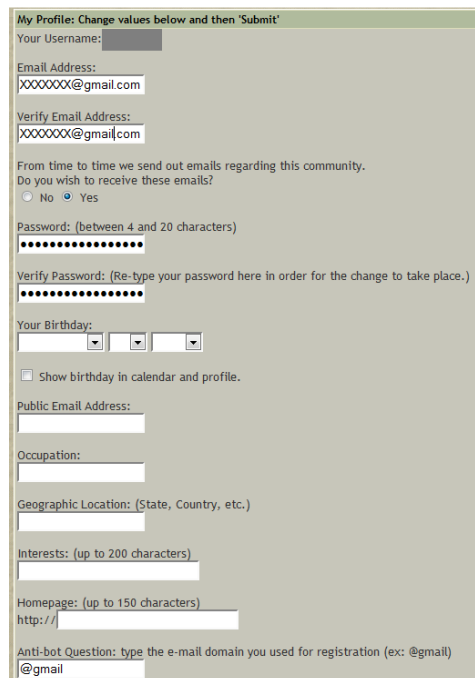
4.1 Εισαγωγή

Η πρώτη μορφή CAPTCHA που παρουσιάστηκε ήταν απλές ερωτήσεις στις οποίες ο χρήστης έπρεπε να απαντήσει εισάγοντας την απάντηση με κείμενο. Αυτή η μέθοδος αποδείχθηκε σύντομα ανεπαρκής και οδήγησε σε τεστ που βασίζονται στην παραμόρφωση κειμένου, η λύση των οποίων ήταν αρκετά πιο δύσκολη. Καθώς αυτή βασίζεται στην οπτική αναγνώριση των χαρακτήρων, σε αρκετές εφαρμογές παρέχεται και η δυνατότητα αναγνώρισης φωνής για τους χρήστες με προβλήματα όρασης. Τέλος, τα τελευταία χρόνια όπου άρχισαν να πληθαίνουν οι διαμαρτυρίες για την ευχρηστία και την αποτελεσματικότητα των παραπάνω μεθόδων έχουν παρουσιαστεί εφαρμογές που βασίζονται στην αναγνώριση εικόνας ή και την εκτέλεση κάποιας συγκεκριμένης ενέργειας από τον χρήστη που οι μηχανές είναι δύσκολο να επιτελέσουν.

Σε αυτό το κεφάλαιο θα μελετήσουμε τα διάφορα είδη τεστ που έχουν αναπτυχθεί και θα δούμε τα κύρια χαρακτηριστικά τους, τα πλεονεκτήματα που έχουν καθώς και τα προβλήματα που δημιουργούν. Τέλος θα αναφέρουμε και τις τεχνικές ανάλυσης της συμπεριφοράς του χρήστη, η οποία αν και δεν είναι άμεσα ένα τεστ CAPTCHA επιτελεί παρόμοιο σκοπό.

4.2 Ερωτήσεις κειμένου

Η πρώτη μορφή CAPTCHA που εμφανίστηκε πριν από την επίσημη μελέτη του προβλήματος ήταν απλές ερωτήσεις κειμένου μέσα σε μία φόρμα που ο χρήστης έπρεπε να απαντήσει. Οι ερωτήσεις βασίζονταν στην λογική και η απάντηση τους ήταν μία απλή λέξη.



The image shows a web form titled "My Profile: Change values below and then 'Submit'". The form contains several input fields and a CAPTCHA question. The fields are: "Your Username:" (text input), "Email Address:" (text input with value "XXXXXXXX@gmail.com"), "Verify Email Address:" (text input with value "XXXXXXXX@gmail.com"), "Password:" (password input with value "*****"), "Verify Password:" (password input with value "*****"), "Your Birthday:" (three dropdown menus), "Public Email Address:" (text input), "Occupation:" (text input), "Geographic Location: (State, Country, etc.)" (text input), "Interests: (up to 200 characters)" (text input), "Homepage: (up to 150 characters)" (text input with value "http://"), and "Anti-bot Question: type the e-mail domain you used for registration (ex: @gmail)" (text input with value "@gmail"). There are also radio buttons for "Do you wish to receive these emails?" (No/Yes) and a checkbox for "Show birthday in calendar and profile."

Εικόνα 4-1 Η φόρμα εγγραφής και εισαγωγής στοιχείων της εφαρμογής δημιουργίας forum phpBB. Στο τέλος φαίνεται μία απλή ερώτηση αναγνώρισης των ανθρώπων – χρηστών, όπου ζητείται το domain της υπηρεσίας email που έχουν εισάγει.

Πρόκειται για απλές ερωτήσεις κειμένου, που έχουν προφανής απαντήσεις, όπως για παράδειγμα ποιο είναι το όνομα του site, ένα συγκεκριμένο γράμμα από κάποιο από τα στοιχεία του χρήστη, το αποτέλεσμα μίας αριθμητικής πράξης με γράμματα αντί για νούμερα κ.ο.κ. Πολλές φορές ο χρήστης καλούνταν να επιλέξει ή να αποεπιλέξει ένα πεδίο checkbox, ώστε να δηλώσει ότι είναι άνθρωπος ή ότι δεν είναι μηχανή.

Το γενικότερο χαρακτηριστικό των παραπάνω είναι ότι το πλήθος των ερωτήσεων είναι αρκετά μικρό, με πολλές σελίδες να έχουν μόνο μία ερώτηση. Η άμυνα απέναντι στα αυτοματοποιημένα εργαλεία επίθεσης δεν έγκειται στην δυσκολία επίλυσης του προβλήματος, το οποίο πολλές φορές είναι η ίδια λέξη για οποιαδήποτε φόρμα, αλλά στο ότι καθώς αυτά τα εργαλεία επιτίθενται σε εκατοντάδες σελίδες καθημερινά δεν θα ασχοληθεί κάποιος άνθρωπος με το να εισάγει την κατάλληλη λύση στο εργαλείο επίθεσης.

Αν και η παραπάνω λύση μπορεί να φέρει κάποια αποτελέσματα είναι πολύ εύκολο να ξεπεραστεί με την ανθρώπινη παρέμβαση, καθώς ένας άνθρωπος μπορεί να δει την ερώτηση ή τις ερωτήσεις και να εισάγει τις απαντήσεις. Επιπλέον η βασική λογική ότι ένας επιτιθέμενος δεν πρόκειται να ασχοληθεί με ένα συγκεκριμένο site σημαίνει ότι η λύση μπορεί να χρησιμοποιηθεί σε σελίδες με πολύ χαμηλή κυκλοφορία που δεν θα τραβήξουν την προσοχή, ενώ η δημιουργία μίας βιβλιοθήκης που θα χρησιμοποιηθεί σε πολλές σελίδες δεν είναι δυνατή.

Text CAPTCHA Examples

This page shows examples of the logic questions returned by the [web service API](#).

- | Which of tooth, hair or Richard is the name of a person?
- | The number of body parts in the list chest, thumb and mosquito is?
- | What is the 6th digit in 3888669?
- | The 4th letter in "refuses" is?
- | Which of cat, nose, apple or cow is part of a person?
- | The word "mailbag" has how many letters?
- | Vocalizing, forceful, postseason: the word starting with "f" is?
- | Which of toe, stomach, tongue or heart is below the waist?

Εικόνα 4-2 Κάποιες τυπικές ερωτήσεις βασισμένες σε κείμενο, όπως φαίνονται από την βιβλιοθήκη text CAPTCHA.

Η σύγχρονη μορφή αυτό του είδους CAPTCHA είναι η συγκέντρωση μεγάλου αριθμού ερωτήσεων σε μία κεντρική βάση δεδομένων και η χρήση τους από πολλές εφαρμογές ταυτόχρονα. Καθώς το πλήθος των ερωτήσεων πλησιάζει στα εκατομμύρια δεν υπάρχει μία βασική μορφή ώστε να λύνεται με έναν γενικό αλγόριθμο, αλλά υπάρχουν χιλιάδες βασικά είδη ώστε δεν είναι απλό να αναγνωριστούν από ένα αυτόματο εργαλείο για να λυθούν.

4.2.1 Μειονεκτήματα

4.2.1.1 Έλλειψη πολυγλωσσικής υποστήριξης

Τα CAPTCHA που βασίζονται σε ερωτήσεις κειμένου έχουν ένα βασικό μειονέκτημα σε σχέση με τα υπόλοιπα: Σε όλες τις έτοιμες βιβλιοθήκες τέτοιου είδους οι ερωτήσεις είναι στα αγγλικά. Αυτό τις κάνει προσβάσιμες μόνο στους χρήστες που ξέρουν αγγλικά ώστε να μπορούν να καταλάβουν το νόημά τους. Αξίζει να παρατηρήσουμε πώς το γεγονός πως ένας χρήστης χρησιμοποιεί μία αγγλική σελίδα δεν σημαίνει πως γνωρίζει και την γλώσσα. Αρκετοί χρήστες έχουν ελάχιστες γνώσεις της γλώσσας πέρα από το τι ζητάνε τα βασικά πεδία μίας φόρμας. Κατά συνέπεια μία τέτοια ερώτηση θα κάνει την χρήση τέτοιων σελίδων απαγορευτικές για αυτούς.

Αν και μπορούν να γραφούν εφαρμογές που περιέχουν τέτοιες ερωτήσεις σε άλλες γλώσσες υπάρχουν δύο μεγάλα εμπόδια. Το πρώτο είναι πως η εργασία που απαιτείται για την προσθήκη μίας νέας γλώσσας πολλαπλασιάζεται ανάλογα με το πλήθος τους, αφού κάθε μία ερώτηση πρέπει να εισάγεται σε όλες τις υποστηριζόμενες γλώσσες.

4.2.1.2 Αναγκαιότητα μεγάλου όγκου ερωτήσεων

Το μεγαλύτερο όμως πρόβλημα είναι πως η ανάπτυξη μίας νέας εφαρμογής δεν είναι σίγουρο ότι θα αποφέρει το επιθυμητό αποτέλεσμα, εκτός και αν περιέχει αρκετές χιλιάδες ερωτήσεις. Όπως αναφέραμε προηγουμένως σε αντίθετη περίπτωση ένας επιτιθέμενος θελήσει μπορεί να εισάγει χειροκίνητα τις απαντήσεις για κάθε ερώτηση ή αυτοματοποιημένα για κάποιους τύπους κοινών ερωτήσεων. Καθώς οι ερωτήσεις είναι προκαθορισμένες ένα αυτοματοποιημένο εργαλείο μπορεί μετά να χρησιμοποιεί τις αποθηκευμένες λύσεις.

4.2.2 Τυπικές επιθέσεις

4.2.2.1 Ευκολία μη αυτοματοποιημένων επιθέσεων

Η βασική επίθεση που μπορεί να γίνει σε ένα τέτοιου είδους captcha είναι η χειροκίνητη εισαγωγή των λύσεων ή κάποιων βασικών κανόνων για την επίλυση των παρόμοιων ερωτήσεων. Το πόσο πρακτική είναι αυτή η λύση εξαρτάται από την ποικιλία των ερωτήσεων και την πολυπλοκότητά τους, αλλά καθώς οι δημιουργοί τέτοιων σελίδων επιδιώκουν να μην δυσκολέψουν τους χρήστες η δυσκολία τους είναι ιδιαίτερα μικρή.

4.2.2.2 Χρήση αλγορίθμων τεχνητής νοημοσύνης

Στην περίπτωση που υπάρχουν εκατομμύρια ερωτήσεις η παραπάνω επίθεση δεν έχει νόημα. Σε τέτοιες βιβλιοθήκες υπάρχουν συγκεκριμένα είδη ερωτήσεων που με πολλές παραλλαγές, πχ η εφαρμογή μπορεί να ζητάει ένα συγκεκριμένο γράμμα μίας λέξης ή την αναγνώριση μίας συγκεκριμένης λέξης από μία ομάδα (πχ να παραθέτει κάποια φρούτα και να ρωτάει ποιο έχει ένα συγκεκριμένο χρώμα). Ο τρόπος που έχει επιχειρηθεί για να σπάσουν τέτοιου είδους τεστ είναι η χρήση νευρωνικών δικτύων με τεχνικές υπολογιστικής μάθησης.

Βέβαια η χρήση εφαρμογών τεχνητής νοημοσύνης παρουσιάζει προβλήματα για τους επιτιθέμενους. Το πρώτο από αυτά είναι πως πρόκειται για αρκετά πολύπλοκες εφαρμογές που δεν έχουν αρκετά μεγάλα ποσοστά επιτυχίας, ενώ ακόμα μεγαλύτερο

πρόβλημα είναι η μεγάλες απαιτήσεις που έχουν τέτοια προγράμματα σε πόρους, δηλαδή σε μνήμη και υπολογιστική ισχύ.

Μία πρόσφατη έρευνα επέδειξε την χρήση της μηχανής αναζήτησης Wolfram Alpha για την επίθεση στις ερωτήσεις του textCaptcha, μίας βιβλιοθήκης με ερωτήσεις κειμένου. Η μηχανή αναζήτησης Wolfram Alpha χρησιμοποιεί αλγορίθμους τεχνητής νοημοσύνης για να απαντήσει σε ερωτήσεις που έχουν διατυπωθεί στην ανθρώπινη γλώσσα, με σκοπό να παρουσιάσει μία αναλυτική απάντηση για το νόημα της ερώτησης. Η πρωτοτυπία της συγκεκριμένης επίθεσης είναι ότι δεν χρησιμοποιεί πόρους του επιτιθέμενου αλλά κάνει χρήση μίας δωρεάν υπηρεσίας που είναι διαθέσιμη στο διαδίκτυο. Επίσης δείχνει ότι πολύ σύντομα οι συγκεκριμένες τεχνικές θα έχουν εξελιχθεί σε ικανοποιητικό επίπεδο ώστε να μπορούν να ξεπεράσουν τέτοιου είδους ερωτήσεις.



Εικόνα 4-3 Η μηχανή αναζήτησης Wolfram Alpha δίνει μία επιτυχή απάντηση σε μία ερώτηση ενός τεστ CAPTCHA

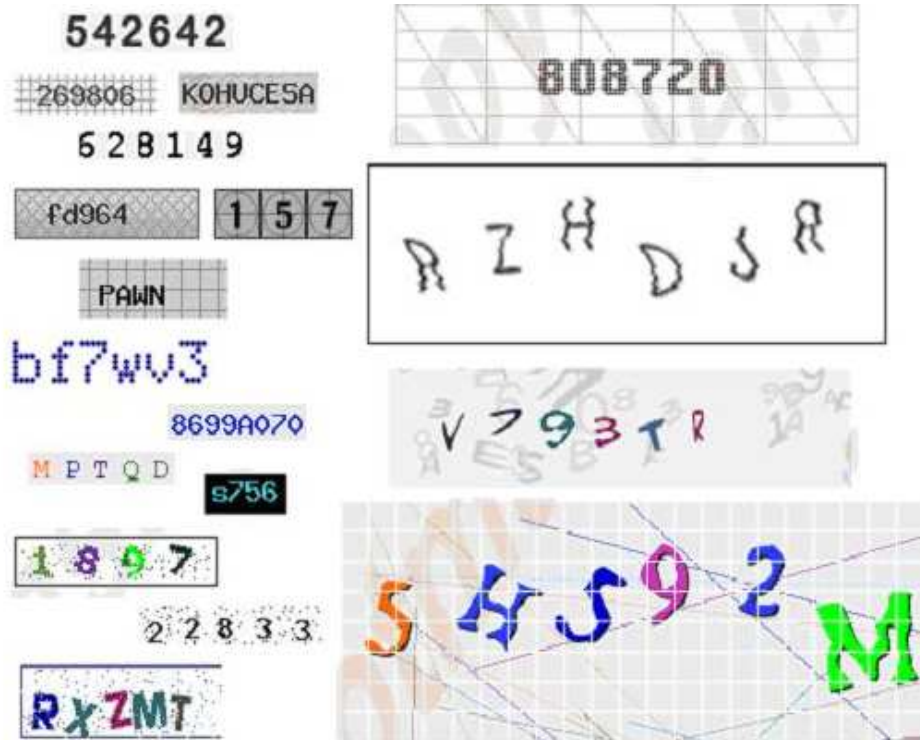
4.2.2.3 Χρήση λογικών εκφράσεων

Σήμερα όμως οι περισσότερες επιθέσεις γίνονται με την δημιουργία κανόνων για την επίλυση συγκεκριμένων ειδών ερωτήσεων, συνήθως με την χρήση λογικών εκφράσεων. Αυτή η επίθεση αντιμετωπίζεται με την χρήση ακόμη περισσότερων ειδών ερωτήσεων ώστε οι επιτυχείς προσπάθειες επίλυσης να είναι ένα πολύ μικρό ποσοστό από τις συνολικές προσπάθειες.

4.3 CAPTCHA παραμόρφωσης γραμμάτων

4.3.1 Περιγραφή

Το πρώτο πραγματικό είδος CAPTCHA που αναπτύχθηκε ήταν εικόνες οι οποίες περιέχουν κάποιες λέξεις των οποίων τα γράμματα είναι παραμορφωμένα. Η παραμόρφωση επιτυγχάνεται με την περιστροφή των χαρακτήρων, την εναλλαγή του χρώματός τους και τη σχεδίαση τυχαίων γραμμών στην εικόνα με τέτοιο τρόπο ώστε να μην είναι εύκολη η ανάγνωσή τους.



Εικόνα 4-4 Παραδείγματα CAPTCHA αναγνώρισης γραμμάτων.

Πρόκειται για το πιο διαδεδομένο είδος και κατά συνέπεια έχει γνωρίσει και την μεγαλύτερη ανάπτυξη. Αν και αρχικά η παραμόρφωση ήταν μικρής έκτασης με το πέρασμα του χρόνου έχουν γίνει αρκετές έρευνες με σκοπό την αύξηση της δυσκολίας ανάγνωσης από τους υπολογιστές.



Εικόνα 4-5 Ένα πρωτότυπο CAPTCHA αναγνώρισης γραμμάτων, το οποίο αφού παρουσιάζει του πρόβλημα λέει στους χρήστες να το αγνοήσουν. Παρόλα αυτά είναι πιθανό πολλοί χρήστες να συμπληρώσουν το πεδίο από συνήθεια χωρίς να διαβάσουν τις οδηγίες

4.3.2 Μειονεκτήματα

4.3.2.1 Δυσκολία χρήσης

Η παραμόρφωση χαρακτήρων έχει κάποια βασικά μειονεκτήματα. Το μεγαλύτερο είναι ότι στην προσπάθεια να γίνουν οι χαρακτήρες όσο το δυνατόν πιο δυσανάγνωστοι από τους υπολογιστές καταλήγουν να είναι και αρκετά δύσκολο να διαβαστούν από τους ανθρώπους. Για παράδειγμα η εισαγωγή τυχαίων γραμμών πάνω από τους χαρακτήρες μπορεί να καλύψει τις διαφορές που έχουν κάποια παρόμοια γράμματα, όπως το κεφαλαίο L και το πεζό l. Επίσης η τυχαία εναλλαγή χρωμάτων μπορεί να κάνει κάποιο γράμμα να μην

ξεχωρίζει από το φόντο της εικόνας, ενώ η περιστροφή είναι πιθανόν να οδηγήσει στη σύγκριση κάποιων χαρακτήρων, όπως το 6 και το 9 ή το μηδέν και το κεφαλαίο 0.

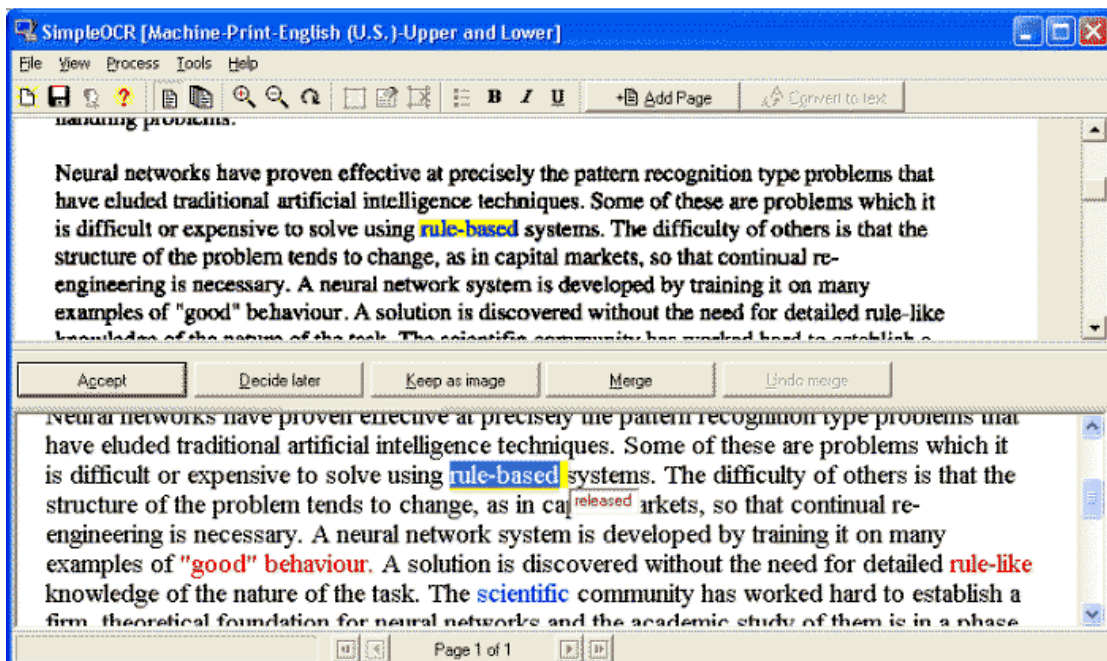
4.3.2.2 Αποκλεισμός χρηστών με ειδικές ανάγκες

Τα CAPTCHA αυτού του είδους θεωρούν ότι ο χρήστης που θα κληθεί να τα συμπληρώσει μπορεί να δει την εικόνα, αποκλείοντας τα άτομα με ειδικές ανάγκες που έχουν προβλήματα όρασης.

4.3.3 Τυπικές επιθέσεις

4.3.3.1 Οπτική αναγνώριση χαρακτήρων

Ο κλάδος της επιστήμης των υπολογιστών που ασχολείται με το πρόβλημα που θέτουν τα CAPTCHA παραμόρφωσης γραμμάτων είναι η ανάλυση εικόνας και η οπτική αναγνώριση χαρακτήρων. Για την υπολογιστική επίλυση της ερώτησης χρειάζεται ο επιτιθέμενος να επεξεργαστεί την εικόνα με έναν τέτοιο αλγόριθμο ώστε να μπορέσει να διαχωρίσει τα γράμματα και να αναγνωρίσει το κάθε ένα από αυτά.



Εικόνα 4-6 Ένα πρόγραμμα OCR ενώ προσπαθεί να διαβάσει ένα βιβλίο.

Αυτοί οι αλγόριθμοι έχουν πολύ μεγάλο υπολογιστικό κόστος και απαιτούν αρκετό χρόνο. Επίσης καθώς πρόκειται για έναν συγκεκριμένο ερευνητικό κλάδο με αρκετό δημοσιευμένο υλικό τα μειονεκτήματα κάθε αλγορίθμου είναι γνωστά και πλέον τα σύγχρονα CAPTCHA σχεδιάζονται με σκοπό να τα εκμεταλλευτούν, ώστε να μειώσουν το ποσοστό επιτυχίας. Όμως με το πέρασμα του χρόνου ο κλάδος ωριμάζει και αναπτύσσονται πιο αποτελεσματικοί αλγόριθμοι, ενώ οι επιδόσεις των υπολογιστών συνεχίζουν να αυξάνονται εκθετικά. Όμως από την ερευνητική εφαρμογή ενός αλγορίθμου μέχρι την πρακτική χρήση του υπάρχει μία μεγάλη απόσταση και χρειάζονται αρκετές τεχνικές γνώσεις.

4.3.3.2 Προώθηση σε τρίτους

Ένας τρόπος αποφυγής του προβλήματος είναι η προώθηση του σε άλλους ανθρώπους. Αυτό μπορεί να γίνει με δύο τρόπους.

Ο ένας είναι η προώθηση του CAPTCHA σε μία άλλη ιστοσελίδα που ανήκει στον επιτιθέμενο. Αρχικά ο επιτιθέμενος παίρνει την φόρμα από την σελίδα – στόχο και απομονώνει την εικόνα με το CAPTCHA. Έπειτα χρησιμοποιεί την ίδια εικόνα σε μία φόρμα στην δική του σελίδα σαν CAPTCHA. Όταν ο χρήστης την συμπληρώσει χρησιμοποιεί την απάντησή του για να απαντήσει στο αρχικό τεστ. Όμως για να είναι επιτυχής αυτή η τεχνική προϋποθέτει ότι η σελίδα του επιτιθέμενου έχει αρκετούς χρήστες και ότι όταν πάρει μία απάντηση θα προλάβει να την δώσει μέσα σε έγκυρο χρονικό διάστημα στην αρχική φόρμα, καθώς αν το CAPTCHA είναι σωστά υλοποιημένο θα έχει ένα μέγιστο χρόνο ζωής.

Ένας άλλος τρόπος προώθηση του προβλήματος είναι η χρήση εταιρειών που ειδικεύονται στην αναγνώριση CAPTCHA, κυρίως σε χώρες της ανατολικής Ασίας. Αυτές έχουν υπαλλήλους που η δουλειά τους είναι να λύνουν CAPTCHA, και συνήθως αρκετά χαμηλά ποσά για την αναγνώριση χιλιάδων CAPTCHA. Πρόκειται για μία τακτική που δεν μπορεί να αντιμετωπιστεί, αλλά καθώς χρειάζεται την καταβολή ενός χρηματικού ποσού από τον επιτιθέμενο δεν είναι τόσο διαδεδομένη.

4.4 CAPTCHA Ήχου

4.4.1 Περιγραφή

Μετά την εμφάνιση των πρώτων CAPTCHA παραμόρφωσης γραμμάτων έγινε σύντομα αντιληπτό ότι ένα από τα μεγαλύτερα μειονεκτήματα τους, όπως έχουμε ήδη αναφέρει, είναι η αδυναμία επίλυσής τους από άτομα με προβλήματα όρασης. Επιπλέον η τάση ψηφιοποίησης παλιότερων εντύπων με σκοπό την διατήρησή τους και την ευκολότερη ηλεκτρονική διανομή τους οδήγησε σε μία αλματώδη βελτίωση των τεχνικών οπτικής αναγνώρισης χαρακτήρων.



Εικόνα 4-7 Ένα τυπικό CAPTCHA ήχου

Οι παραπάνω λόγοι οδήγησαν στην εμφάνιση των CAPTCHA ήχου. Πρόκειται για τεστ στα οποία ο χρήστης ακούει ένα συγκεκριμένο ηχητικό απόσπασμα και πρέπει να γράψει το περιεχόμενό του, το οποίο συνήθως είναι κάποια γράμματα ή λέξεις.

4.4.2 Μειονεκτήματα

4.4.2.1 Ανάγκη καλής γνώσης της γλώσσας

Αν και η αρχική αιτία χρήσης των ακουστικών CAPTCHA ήταν η επίλυση των προβλημάτων πρόσβασης ατόμων με ειδικές ανάγκες, τελικά κατέληξε να δημιουργεί περισσότερα προβλήματα από όσα επιλύει. Το κυριότερο από αυτά είναι η δυσκολία ή η αδυναμία επίλυσης τους από χρήστες που δεν μιλούν την γλώσσα στην οποία είναι ηχογραφημένο το τεστ.

Σε αντίθεση με τα προηγούμενα CAPTCHA παραμόρφωσης γραμμάτων και κειμένου, τα οποία απαιτούν από τον χρήστη να μπορεί να αναγνωρίσει συγκεκριμένους χαρακτήρες ή να κατανοήσει το νόημα απλών προτάσεων, οι ηχητικές δοκιμασίες απαιτούν από τον χρήστη να έχει αρκετά μεγαλύτερη εξοικείωση με την γλώσσα, αφού η αναγνώριση του ομιλούντος λόγου είναι αρκετά δυσκολότερη. Με την πλειοψηφία των CAPTCHA αυτού του είδους να είναι στα αγγλικά, οι μη αγγλόφωνες πρακτικά αποκλείονται από την χρήση τους.

4.4.2.2 Έλλειψη κοινού τεχνικού προτύπου

Ένα αρκετά μεγάλο τεχνικό πρόβλημα αυτών των δοκιμασιών είναι το γεγονός ότι δεν υπάρχει κάποια κοινώς αποδεκτή προδιαγραφή στην γλώσσα HTML, στην οποία είναι γραμμένες οι ιστοσελίδες, για την αναπαραγωγή ενός αρχείου ήχου. Κατά συνέπεια είναι απαραίτητη η χρήση διαφορετικών τεχνικών για κάθε είδος και έκδοση προγράμματος περιήγησης, πράγμα που ανεβάζει το κόστος και την δυσκολία ανάπτυξης της ιστοσελίδας.



Εικόνα 4-8 Τυπικά τα ακουστικά CAPTCHA προσφέρουν στον χρήστη την δυνατότητα να κατεβάσει το αρχείο ήχου για την περίπτωση που ο φυλλομετρητής του δεν μπορεί να το αναπαράγει στην σελίδα.

Οι περισσότερες σελίδες παρακάμπτουν αυτό το πρόβλημα χρησιμοποιώντας ένα πρόγραμμα γραμμένο στην τεχνολογία Flash της εταιρείας Adobe, όμως η εγγενής υποστήριξη του Flash δεν είναι υποχρεωτική από του φυλλομετρητές, απαιτώντας από τους χρήστες να εγκαταστήσουν επιπλέον λογισμικό για την αναπαραγωγή του, όπου και πάλι μπορεί να μην είναι διαθέσιμο για το λειτουργικό σύστημα που χρησιμοποιούν, όπως για παράδειγμα για τα κινητά τηλέφωνα με iOS ή Android όπου η υποστήριξη έχει πλέον σταματήσει.

4.4.2.3 Δυσκολία δημιουργίας

Μία ιδιαιτερότητα που έχουν οι δοκιμασίες ήχου είναι ότι η επεξεργασία ήχου είναι αρκετά απαιτητική σε υπολογιστικούς πόρους, πολύ περισσότερο από την επεξεργασία εικόνων. Υπάρχουν δύο τεχνικές που μπορούν να χρησιμοποιηθούν, η χρήση ήδη ηχογραφημένων μηνυμάτων και η τεχνητή σύνθεσή τους από κατάλληλο λογισμικό.

Η χρήση έτοιμων μηνυμάτων δεν είναι ιδιαίτερα πρακτική, καθώς πρόκειται για μία χρονοβόρα και ακριβή διαδικασία. Επιπλέον προϋποθέτει πως θα υπάρχει ένα πεπερασμένο πλήθος μηνυμάτων, τα οποία μπορούν να λυθούν μία φορά με μη αυτοματοποιημένο τρόπο και να χρησιμοποιούνται.

Η τεχνητή σύνθεση του ήχου, η οποία μπορεί να γίνει με χρήση λογισμικού σύνθεσης φωνής (Computer Generated Speech Software) έχει με τη σειρά της το μειονέκτημα ότι η παραγόμενη φωνή δεν έχει την χροιά και τυχόν προφορές και ιδιαιτερότητες που έχει η φωνή ενός κανονικού ανθρώπου, καθιστώντας της επίλυση αρκετά πιο εύκολη.

Συνήθως ακολουθείται μία μέση οδός, με την χρήση κάποιων ήδη ηχογραφημένων μηνυμάτων και την ανάμειξή τους με επιπλέον θορύβους περιβάλλοντος να εισάγονται στο παραγόμενο αποτέλεσμα ώστε να ανέβει το επίπεδο δυσκολίας. Αυτοί οι υπολογισμοί έχουν το μειονέκτημα ότι είναι αρκετά απαιτητική και αυξάνουν σημαντικά το φορτίο του διακομιστή που τους αναλαμβάνει.

4.4.2.4 Αποκλεισμός χρηστών με ειδικές ανάγκες

Είναι προφανές ότι τα CAPTCHA ήχου αποκλείουν τους χρήστες με προβλήματα ακοής από την χρήση τους, καθώς δεν μπορούν να ακούσουν το περιεχόμενο του τεστ.

4.4.3 Τυπικές επιθέσεις

4.4.3.1 Ακουστικό αποτύπωμα

Ένα ακουστικό αποτύπωμα είναι μία συμπυκνωμένη ψηφιακή αναπαράσταση ενός ηχητικού αποσπάσματος, το οποίο μπορεί να χρησιμοποιηθεί για την μετέπειτα αναγνώρισή του. Τα πρώτα ηχητικά CAPTCHA αποτελούνταν συνήθως από 26 αποσπάσματα ήχου, ένα για κάθε γράμμα του αλφαβήτου, με την ίδια τη δοκιμασία να αποστέλλει τυχαία κάποια από αυτά και να ζητάει από τον χρήστη να τα αναγνωρίσει.

Με την χρήση των ακουστικών αποτυπωμάτων η επίθεση σε τέτοιου είδους δοκιμασίες είναι τετριμμένη, καθώς αρκεί η δημιουργία ενός αποτυπώματος για κάθε γράμμα και έπειτα η σύγκρισή τους με το μήνυμα της δοκιμασίας. Για την αντιμετώπιση αυτής της επίθεσης αρχικά χρησιμοποιήθηκε η εισαγωγή τυχαίων θορύβων στο μήνυμα, με σκοπό την παραμόρφωση του αποτυπώματος. Δυστυχώς η τεχνολογία έχει μία εγγενή τάση να απομονώνει τέτοιους θορύβους και να μην τους δίνει μεγάλη σημασία.

Πλέον η εισαγωγή ήχων γίνεται με ειδικά μελετημένο τρόπο ώστε να παραμορφώνεται το παραγόμενο αποτύπωμα ώστε να μπερδεύονται οι αντίστοιχοι αλγόριθμοί, χωρίς να μπερδεύεται ο χρήστης, κάτι το οποίο δεν είναι ιδιαίτερα εύκολο. Ποιο πρόσφορη είναι η χρήση λέξεων και προτάσεων με τυχαία σειρά και χροιά, το πλήθος των οποίων δεν επιτρέπει την δημιουργία χωριστών αποτυπωμάτων για κάθε μία από αυτές.

4.4.3.2 Αυτόματη αναγνώριση ομιλίας

Σε αναλογία με την οπτική αναγνώριση χαρακτήρων, υπάρχουν αλγόριθμοι αναγνώρισης ομιλίας (Automated Speech Recognition - ASR) οι οποίοι μετατρέπουν ένα ακουστικό απόσπασμα σε κείμενο. Αν και όταν είχαν αρχικά παρουσιαστεί τα ηχητικά CAPTCHA η τεχνολογία βρισκόταν στα βρεφικά της στάδια, πλέον έχει αναπτυχθεί αρκετά ώστε να μπορεί να αναγνωρίζει ήχους με αρκετή ακρίβεια.

Για να αποτρέψουν την εύκολη επίλυση των δοκιμασιών από αυτοματοποιημένα προγράμματα οι δημιουργοί τους τα κάνουν πιο δύσκολα εισάγοντας επιπλέον θορύβους, άσχετους με το περιεχόμενο ή τοπικές προφορές, με σκοπό να μπερδέψουν τα αντίστοιχα προγράμματα. Όμως αυτό έχει οδηγήσει όπως και στην περίπτωση των CAPTCHA παραμόρφωσης γραμμάτων σε έναν αγώνα δρόμου καθώς και οι αντίστοιχοι αλγόριθμοι βελτιώνονται συνεχώς.

Ένα μειονέκτημα αυτής της επίθεσης είναι ότι είναι αρκετά απαιτητική σε υπολογιστικούς πόρους, περισσότερο από την οπτική αναγνώριση χαρακτήρων, καθώς ένα ηχητικό απόσπασμα συνήθως καταλαμβάνει περισσότερους πόρους και απαιτεί αρκετή υπολογιστική ισχύ για την επεξεργασία του, υποχρεώνοντας τον επιτιθέμενο να έχει στη διάθεσή του ιδιαίτερα εκτενής υποδομές.

4.4.3.3 Προώθηση σε τρίτους

Όπως και σε όλα τα παραπάνω τεστ, τα CAPTCHA μπορούν να προωθηθούν σε ανθρώπους ώστε να τα λύσουν και να στείλουν την σωστή απάντηση. Παρόλο που δεν υπάρχει επιτυχή άμυνα σε αυτού του είδους επίθεση οι ηχητικές διαδικασίες παρουσιάζουν μία ιδιαιτερότητα σε σχέση με τα υπόλοιπα τεστ.

Συγκεκριμένα απαιτούν από τον χρήστη να έχει αρκετά καλή γνώση της γλώσσας στην οποία βρίσκεται το κείμενο. Συνήθως τα τεστ είναι στα αγγλικά και οι οργανισμοί οι οποίοι αναλαμβάνουν την αυτόματη επίλυση των τεστ βρίσκονται σε χώρες που οι χρήστες δεν χρησιμοποιούν αγγλικά σαν μητρική γλώσσα, καθιστώντας ποιο δύσκολη την εύρεση υπαλλήλων που να μπορούν να τα επιλύσουν.

4.5 CAPTCHA Αναγνώρισης εικόνας

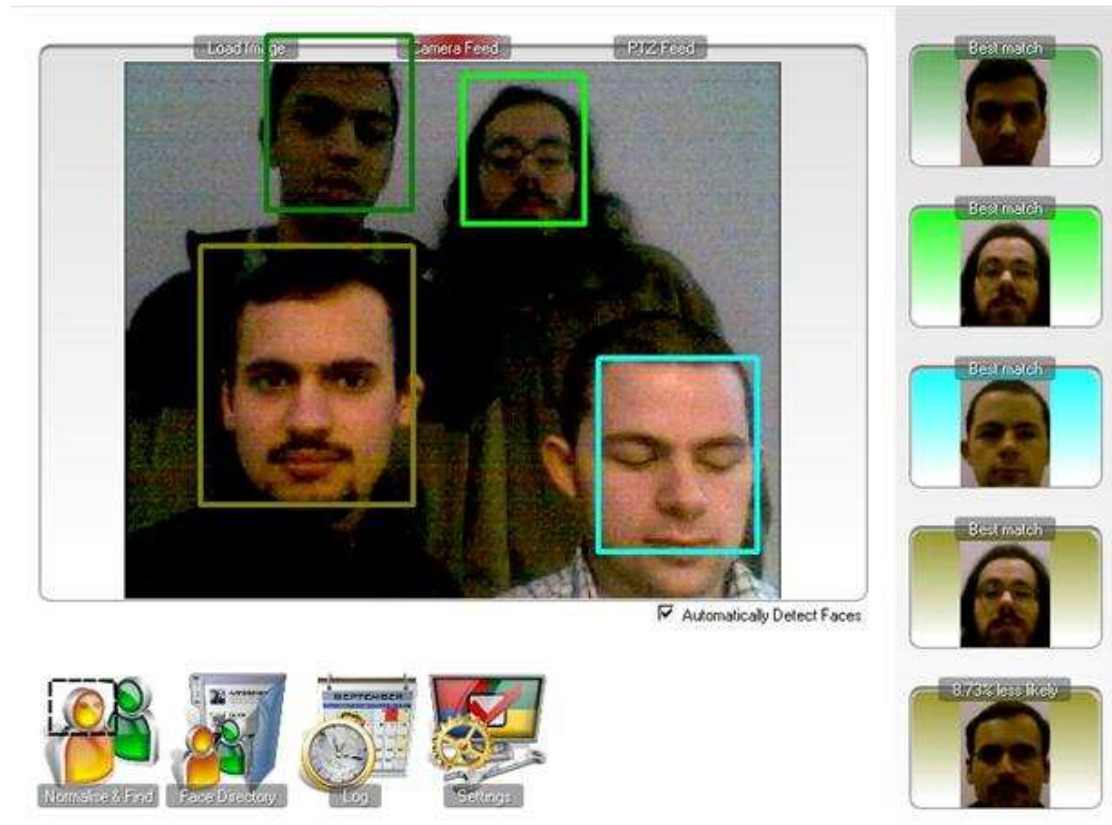
4.5.1 Περιγραφή

Τα CAPTCHA αναγνώρισης εικόνας ζητούν συνήθως από τους χρήστες να αναγνωρίσουν ένα συγκεκριμένο αντικείμενο από ένα πλήθος δοθέντων εικόνων. Σε σχέση με τα προηγούμενα είναι το ποιο απλό στην χρήση, καθώς η διαδικασία επίλυσης είναι αρκετά πιο σύντομη και λιγότερο κουραστική για τους ανθρώπους.



Εικόνα 4-9 Ένα τυπικό CAPTCHA αναγνώρισης εικόνας. Ο χρήστης πρέπει να επιλέξει το αντικείμενο που του ζητείται στην ερώτηση που έχει τεθεί. Για να αποφευχθεί η τυχαία επιλογή μίας εικόνας κατά την προσπάθεια επίλυσης το συγκεκριμένο τεστ θέτει στον χρήστη τέσσερις ερωτήσεις, μειώνοντας δραματικά τις πιθανότητες επίλυσης.

Ένα από τα βασικά προβλήματα που έχουν όλα τα παραπάνω είδη CAPTCHA είναι ότι με το πέρασμα του χρόνου η δυσκολία τους αυξάνεται με σκοπό να δυσκολέψουν τα εργαλεία αυτόματης επίλυσης. Όμως καθώς οι αλγόριθμοι που χρησιμοποιούνται για αυτή ωριμάζουν όλο και περισσότερο η κατάληξη είναι η κατακόρυφη άνοδος της δυσκολίας σε τέτοιο βαθμό που να δυσκολεύουν και τους ανθρώπους που προσπαθούν να τα λύσουν.



Εικόνα 4-10 Ένα πρόγραμμα αναγνώρισης εικόνας το οποίο αναλύει μία εικόνα προσπαθώντας να βρει γνώριμα πρόσωπα. Αν ένα πρόσωπο δεν υπάρχει ήδη στη βάση η διαδικασία είναι αρκετά πιο δύσκολη.

Τα CAPTCHA εικόνων ακολουθούν μια διαφορετική προσέγγιση από την άποψη ότι η αυτόματη επίλυσή τους θα πρέπει να γίνει με την χρήση αλγορίθμων αναγνώρισης προτύπων. Σε αντίθεση με τους τομείς που αναλύσαμε παραπάνω, ο συγκεκριμένος τομέας βρίσκεται ακόμη σε πρώιμο στάδιο ανάπτυξης, από την άποψη ότι δεν υπάρχουν αλγόριθμοι που να έχουν εγγυημένα καλά ποσοστά επιτυχίας, ενώ μπορεί να μπερδευτούν εύκολα από μικρές λεπτομέρειες στις οποίες ένας πραγματικός άνθρωπος δεν θα δώσει καν σημασία.



Εικόνα 4-11 Ένα CAPTCHA αναγνώρισης εικόνας που συνδυάζει και διαφημίσεις.

Επιπλέον οι συγκεκριμένοι αλγόριθμοι είναι άκρως πειραματικοί και δύσκολοι στην χρήση τους, ενώ ακόμη και αν κάποιος έχει τις τεχνικές γνώσεις ώστε να χρησιμοποιήσει έναν από αυτούς θέτουν το επιπλέον εμπόδιο πως είναι αρκετές τάξεις μεγέθους ποιο απαιτητικοί σε υπολογιστικούς πόρους σε σχέση με οποιονδήποτε άλλο αλγόριθμο έχουμε μελετήσει μέχρι τώρα, σε επίπεδο να χρειάζεται τυπικά χρόνος γύρω στο ένα δευτερόλεπτο για την αναγνώριση μίας εικόνας. Αν ένα απλό τεστ δείχνει εννέα εικόνες είναι ξεκάθαρο πως η χρήση μίας τέτοιας λύσης είναι απαγορευτική.

Ένα άλλο σημείο που μπορεί να κάνει την επίλυση τέτοιων τεστ ποιο δύσκολη από αυτόματα εργαλεία χωρίς να δυσκολεύει τους πραγματικούς χρήστες είναι η τροποποίηση των ίδιων των ερωτήσεων. Για παράδειγμα αντί να περιορίζονται σε ερωτήσεις που ζητάνε από τον χρήστη να αναγνωρίσει ένα συγκεκριμένο αντικείμενο μπορούν να ζητάνε να διαλέξει κάποιο είδος, πχ να επιλέξει ένα ζώο από ένα σύνολο εικόνων (ενώ οι υπόλοιπες παρουσιάζουν αντικείμενα).

4.5.2 Μειονεκτήματα

4.5.2.1 Ανάγκη καλής γνώσης της γλώσσας

Αν και από τον χρήστη ζητείται να αναγνωρίσει μία εικόνα, η ερώτηση τίθεται με κείμενο, το οποίο συνήθως είναι στα αγγλικά. Παρόλο που η ερώτηση είναι αρκετά απλή, προϋποθέτει ένα επίπεδο γνώσης της γλώσσας από τον χρήστη ώστε να κατανοήσει το περιεχόμενό της και να επιλέξει την σωστή απάντηση.

4.5.2.2 Μεγάλος όγκος δεδομένων

Για να είναι αποτελεσματικό το τεστ είναι απαραίτητη η ύπαρξη ενός μεγάλου πλήθους εικόνων, με σκοπό να αποκλείεται η μη αυτόματη επίλυση του τεστ. Αυτό σημαίνει ότι χρειάζεται μία αρκετά μεγάλη βιβλιοθήκη εικόνων, αυξάνοντας τον χώρο που καταλαμβάνει το τεστ στον δίσκο του διακομιστή.

Επιπλέον καθώς σε κάθε τεστ χρειάζεται να στέλνεται ένα συγκεκριμένο πλήθος από αυτές τις εικόνες στον χρήστη αυξάνεται και το μέγεθος των δεδομένων που χρειάζεται να κατεβάσει στον υπολογιστή του ώστε να φορτώσει την σελίδα, αφού κάθε εικόνα έχει ένα υπολογίσιμο μέγεθος. Κατά συνέπεια καθυστερεί ο χρόνος φόρτωσης της σελίδας και αυξάνεται ο όγκος δεδομένων που διακινείται μέσω του διακομιστή του ιστοτόπου.

4.5.2.3 Αποκλεισμός χρηστών με ειδικές ανάγκες

Όπως και με τα CAPTCHA παραμόρφωσης χαρακτήρων για να μπορέσει ένας χρήστης να επιλύσει το τεστ πρέπει να μπορεί να δει και τις εικόνες. Έτσι αποκλείονται αυτόματα όλοι οι χρήστες με προβλήματα όρασης.

4.5.3 Τυπικές επιθέσεις

4.5.3.1 Μη αυτόματη αναγνώριση

Σε περίπτωση που το πλήθος των εικόνων που χρησιμοποιούνται σαν πιθανές λύσεις του τεστ είναι σχετικά μικρό είναι δυνατή η αντιστοίχιση κάθε μίας με την λέξη που παρουσιάζει και η χρήση της μετέπειτα με αυτόματο τρόπο. Η μοναδική άμυνα σε αυτή την περίπτωση είναι η ύπαρξη μίας μεγάλης βιβλιοθήκης εικόνων.

4.5.3.2 Προώθηση σε τρίτους

Το μόνο χαρακτηριστικό που έχουν τα CAPTCHA αναγνώρισης εικόνας το οποίο μπορεί να χρησιμοποιηθεί εναντίον τους είναι η ευκολία επίλυσης τους από τους ανθρώπους. Ένα επιτιθέμενος μπορεί να προωθήσει αυτά τα τεστ σε άτομα που πληρώνονται για την επίλυσή τους, τα οποία θα μπορούν να λύσουν μεγάλο αριθμό δοκιμασιών σε μικρό χρονικό διάστημα.

4.6 Διαδραστικά CAPTCHA

4.6.1 Περιγραφή

Τα περισσότερα CAPTCHA υποβάλλουν τον χρήστη σε μία δοκιμασία που συνήθως είναι βαρετή, ενώ η αυξημένη δυσκολία μερικών κάνει την επίλυσή τους ιδιαίτερα δυσάρεστη διαδικασία. Με αυτό το σκεπτικό αναπτύχθηκαν τα διαδραστικά CAPTCHA σκοπεύοντας να μετατρέψουν μία συνήθως βαρετή διαδικασία σε μία ευχάριστη δραστηριότητα.



Εικόνα 4-12 Ένα διαδραστικά CAPTCHA, το οποίο ζητάει από τον χρήστη να σύρει κάποια τμήματα μίας εικόνας στα κατάλληλα σημεία ώστε να την ολοκληρώσει. Το συγκεκριμένο τεστ δεν περιέχει ερώτηση κειμένου.

Σε αυτού του είδους τα τεστ ο χρήστης καλείται να εκτελέσει μία δραστηριότητα με το ποντίκι του για να επιλύσει το CAPTCHA. Συνήθως του δίνεται μία οδηγία μέσω μίας απλής ερώτησης με την οποία του ζητείται να σύρει κάποιες εικόνες ενός συγκεκριμένου αντικειμένου σε συγκεκριμένα σημεία της φόρμας που συμπληρώνει. Αν και η διαδικασία είναι πιο χρονοβόρα σε σχέση με διάφορα άλλα τεστ, η πρωτοτυπία του εγχειρήματος

θεωρείται αρκετή ώστε να αποτρέψει τυχόν αρνητικές εντυπώσεις. Αυτού του είδους τα CAPTCHA δεν έχουν μία συγκεκριμένη μορφή, όπως τα προηγούμενα είδη που μελετήσαμε.



Εικόνα 4-13 Ένα διαφορετικό διαδραστικό CAPTCHA, στο οποίο ο χρήστης είναι απαραίτητο να απαντήσει στην ερώτηση σέρνοντας μία από τις εικόνες στα αριστερά στο πρόσωπο στα δεξιά. Αξίζει να σημειωθεί πως η ερώτηση κειμένου δεν έχει προφανή απάντηση απαιτώντας από τον χρήστη να την κατανοήσει καλά πριν απαντήσει.

Το συγκεκριμένο τεστ παρουσιάζει δύο προκλήσεις στους επίδοξους επιτιθέμενους. Η πρώτη είναι η εύρεση και η κατανόηση της ερώτησης, στην περίπτωση που τίθεται μία τέτοια, διαδικασία που δεν είναι πάντα τετριμμένη. Η δεύτερη είναι πως καθώς οι κινήσεις που γίνονται δεν μπορούν να αντιστοιχηθούν άμεσα στη συμπλήρωση του πεδίου μίας φόρμας χρειάζεται μεγαλύτερος χρόνος και κόπος για να την ανάλυση του τρόπου αποστολής των λύσεων από τους επιτιθέμενους.

Επιπλέον, καθώς η πλειοψηφία των χρηστών έχουν εξοικειωθεί με τα προηγούμενα είδη CAPTCHA, το γεγονός πως είναι τόσο διαφορετικά σε σχέση με τα καθιερωμένα μπορεί να κάνει τους χρήστες να θυμάται την συγκεκριμένη σελίδα, εξαιτίας της πρωτοτυπίας αυτής της λύσης.

4.6.2 Μειονεκτήματα

4.6.2.1 Αποκλεισμός ατόμων με ειδικές ανάγκες

Αυτού του είδους οι λύσεις απαιτούν από τον χρήστη να αναγνωρίσει το περιεχόμενο μίας συγκεκριμένης εικόνας, αποκλείοντας αυτόματα τα άτομα με προβλήματα όρασης.

Επιπλέον η διαδικασία του να σύρει κάποιος μία εικόνα σε έναν συγκεκριμένο στόχο είναι αρκετά απαιτητική αποκλείοντας πέρα από τα άτομα με κινητικά προβλήματα και άλλες ομάδες ατόμων, όπως τους ηλικιωμένους ή αυτούς που δεν έχουν ιδιαίτερη εξοικείωση με την χρήση ηλεκτρονικών υπολογιστών.

4.6.2.2 Δυσκολία χρήσης

Η μορφή των δοκιμασιών αυτού του είδους έχει σαν κοινό χαρακτηριστικό πως παρουσιάζονται σαν γρίφοι, τους οποίους ο χρήστης είναι απαραίτητο να επιλύσει. Κατά συνέπεια χρειάζονται επιπλέον σκέψη σε σχέση με άλλες μορφές, καθώς είναι απαραίτητη η κατανόηση του προβλήματος και της μεθόδου με την οποία πρέπει να λυθεί.

Η πρωτοτυπία αυτών των δοκιμασιών λειτουργεί εναντίων τους στον συγκεκριμένο τομέα, γιατί είναι δυνατόν κάποιος χρήστης να μην μπορεί να καταλάβει ότι πρόκειται για τεστ ή ότι έχει την δυνατότητα να σύρει κάποιες εικόνες.

4.6.2.3 Έλλειψη κοινού τεχνικού προτύπου

Η γλώσσα HTML που περιγράφει την μορφή των ιστοσελίδων δεν δίνει την δυνατότητα υλοποίησης των διαδραστικών τεχνικών που απαιτούνται από τα συγκεκριμένα τεστ. Κατά συνέπεια οι δημιουργοί τους χρειάζεται να καταφύγουν στην χρήση επιπλέον τεχνολογιών, συνηθέστερα είτε στην γλώσσα JavaScript είτε σε κάποια εφαρμογή γραμμένη για την πλατφόρμα Flash. Και οι δύο λύσεις έχουν κάποια εγγενή μειονεκτήματα.

Αν και όλοι οι φυλλομετρητές υποστηρίζουν την γλώσσα JavaScript και μπορούν να εκτελέσουν κώδικα γραμμένο σε αυτή, πρόκειται για μια γλώσσα που δεν παράγει τα ίδια αποτελέσματα σε κάθε έναν από αυτούς. Το επίπεδο της υποστήριξης διαφέρει ανάλογα με την έκδοση κάθε προγράμματος, ενώ μόνο οι πιο σύγχρονες εκδόσεις υποστηρίζουν τις πιο προηγμένες λειτουργίες της. Ακόμη και σε αυτή την περίπτωση τα αποτελέσματα εκτέλεσης ενός προγράμματος μπορεί να διαφέρουν ανάμεσα στους διάφορους φυλλομετρητές.

Αν και τα τελευταία χρόνια γίνεται μεγάλη προσπάθεια να καθιερωθεί ένα κοινό αποδεκτό πρότυπο για τον τρόπο που θα λειτουργεί η γλώσσα, δεν μπορεί να γίνει η παραδοχή πως όλοι οι χρήστες μίας σελίδας που χρησιμοποιεί ένα τεστ τέτοιου είδους θα έχουν την τελευταία έκδοση των κατάλληλων προγραμμάτων. Επιπλέον καθώς συχνά η JavaScript χρησιμοποιείται σαν ένα διάνυσμα επίθεσης για την εγκατάσταση κακόβουλου λογισμικού, πολύ χρήστες απενεργοποιούν την δυνατότητα εκτέλεσης της από τον browser τους, κάνοντας της κατάσταση ακόμη πιο πολύπλοκη.

Η πλατφόρμα ανάπτυξης λογισμικού Flash δεν έχει τέτοιου είδους προβλήματα, καθώς εγγυάται πως μία εφαρμογή θα τρέξει με τον ίδιο τρόπο άσχετα με τον browser που χρησιμοποιεί ο χρήστης. Όμως, όπως αναφέραμε και παραπάνω, απαιτεί από τον χρήστη την εγκατάσταση επιπλέον λογισμικού, το οποίο μπορεί να λείπει από συγκεκριμένα μηχανήματα ενώ δεν είναι καθόλου διαθέσιμο για συγκεκριμένες πλατφόρμες.

4.6.3 Τυπικές επιθέσεις

Σε αντίθεση με τα άλλα είδη CAPTCHA, τα διαδραστικά τεστ δεν έχουν κάποια εγγενή αδυναμία. Ο κυριότερος λόγος είναι ότι συνδυάζουν χαρακτηριστικά από τα CAPTCHA κειμένου, μέσω της ερώτησης με την οποία καθοδηγούν τον χρήστη και από τα CAPTCHA αναγνώρισης εικόνας, καθώς οι ενέργειες που πρέπει να εκτελέσει ο χρήστης βασίζονται στην αναγνώριση των περιεχομένων εικόνων.

Αυτό δεν σημαίνει πως δεν είναι δυνατή η προσπάθεια αυτόματης επίλυσής τους, απλά κανένας από τους ενδεδειγμένους τρόπους δεν είναι ιδιαίτερα αποτελεσματικός. Ακόμη και η προώθηση τους σε τρίτους, παρόλο που θα επιτύχει δεν είναι τόσο αποτελεσματική καθώς είναι ιδιαίτερα χρονοβόρα στην επίλυσή τους.

Το μοναδικό μειονέκτημά τους είναι πως καθώς είναι αρκετά πιο πολύπλοκα στην υλοποίησή τους σε σχέση με τα παραδοσιακά είδη CAPTCHA είναι πιο πιθανό να έχει γίνει κάποιο λάθος σε αυτή που να έχει οδηγήσει στην ύπαρξη κάποιου κενού ασφαλείας, το οποίο μπορεί να εκμεταλλευτεί ο επιτιθέμενος.

4.7 Σύνοψη

Παρακάτω παρουσιάζουμε ένα πίνακα που συνοψίζει τα συμπεράσματα που αναφέραμε παραπάνω.

Τύπος	Πλεονεκτήματα	Μειονεκτήματα	Παρατηρήσεις
Ερωτήσεις κειμένου	<ul style="list-style-type: none"> - Εύκολη υλοποίηση - Μικρός φόρτος στον server - Κατανοητά σε άτομα με ειδικές ανάγκες 	<ul style="list-style-type: none"> - Μικρή αποτελεσματικότητα - Αρκετά δύσκολο να υποστηρίξουν πολλές γλώσσες - Είναι εύκολο να σπάσουν αυτόματα 	Για να είναι πραγματικά αποτελεσματικά χρειάζεται εξαιρετικά μεγάλο πλήθος ερωτήσεων
Παραμόρφωση χαρακτήρων	<ul style="list-style-type: none"> - Μεγάλη αντίσταση στην αυτόματη επίλυση - Σχετικά μικρός φόρτος server 	<ul style="list-style-type: none"> - Πρέπει συνεχώς να γίνονται πιο δύσκολα για να μην λύνονται αυτόματα - Αρκετά δύσκολα στην επίλυσή τους από κανονικούς ανθρώπους - Αδύνατο να επιλυθούν από άτομα με προβλήματα όρασης - Χρονοβόρα στην επίλυσή τους 	<p>Όσο αυξάνεται η αντίστασή τους στις προσπάθειες αυτόματης επίλυσης τόσο πιο δύσκολα γίνονται για τους χρήστες</p> <p>Κάποια στιγμή η τεχνολογία θα ξεπεράσει τις ανθρώπινες δυνατότητες επίλυσης</p>

<p>Ακουστικά</p>	<ul style="list-style-type: none"> - Μπορούν να λυθούν από χρήστες με προβλήματα όρασης 	<ul style="list-style-type: none"> - Αρκετά δύσκολο να υποστηρίξουν πολλές γλώσσες - Δεν είναι εύκολο να αυξηθεί η δυσκολία τους - Σχετικά μεγάλος φόρτος server - Χρονοβόρα στην επίλυσή τους - Βασίζονται σε τεχνολογίες που δεν υποστηρίζονται από όλα τα προγράμματα πειριήγησης 	<p>Αυτή τη στιγμή η τεχνολογία αναγνώρισης φωνής μπορεί σε πολλές περιπτώσεις να έχει καλύτερη απόδοση από τους ανθρώπους</p>
<p>Αναγνώριση εικόνας</p>	<ul style="list-style-type: none"> - Εύκολα στην επίλυση - Αρκετά δύσκολο να επιλυθούν αυτόματα - Μπορούν να χρησιμοποιηθούν σε συνδυασμό με διαφημίσεις προσφέροντας μία επιπλέον πηγή εσόδων 	<ul style="list-style-type: none"> - Αδύνατο να επιλυθούν από άτομα με προβλήματα όρασης - Μεγάλος φόρτος server - Δύσκολα στην υλοποίηση - Είναι αναγκαία η ύπαρξη αρκετά μεγάλου πλήθους εικόνων 	
<p>Τύπος</p>	<p>Πλεονεκτήματα</p>	<p>Μειονεκτήματα</p>	<p>Παρατηρήσεις</p>

<p>Διαδραστικά</p>	<ul style="list-style-type: none"> - Η επίλυση μπορεί να είναι διασκεδαστική για τον χρήστη - Αρκετά δύσκολο να επιλυθούν αυτόματα 	<ul style="list-style-type: none"> - Αδύνατο να επιλυθούν από άτομα με προβλήματα όρασης και κινητικά προβλήματα - Χρειάζεται να δοθεί ιδιαίτερο βάρος στην πολυγλωσσική υποστήριξη - Μεγάλος φόρτος server - Αρκετά δύσκολα στην υλοποίηση - Βασίζονται στη χρήση τεχνολογιών που μπορεί να μην υποστηρίζονται από διάφορα προγράμματα περιήγησης 	
<p>Τύπος</p>	<p>Πλεονεκτήματα</p>	<p>Μειονεκτήματα</p>	<p>Παρατηρήσεις</p>

5 Παρουσίαση εφαρμογής

5.1 Γενικά

Στα πλαίσια της μελέτης των CAPTCHA δημιουργήσαμε μία διαδικτυακή εφαρμογή που χρησιμοποιεί ένα αντιπροσωπευτικό δείγμα που μπορεί να χρησιμοποιήσει ένας web developer και να ενσωματώσει στη σελίδα του. Στόχος είναι η παρουσίαση κάθε λύσης και των δυνατοτήτων που προσφέρει καθώς και η παρουσίαση της διαδικασίας ενσωμάτωσης σε προγραμματιστικό επίπεδο.

Οι περισσότερες λύσεις που υπάρχουν στο διαδίκτυο λειτουργούν σε πολλές πλατφόρμες, καθώς είτε προσφέρουν πολλαπλές υλοποιήσεις για τις πιο δημοφιλείς γλώσσες προγραμματισμού (PHP, Java – JSP, .NET) είτε χρησιμοποιούν Javascript για να είναι ανεξάρτητες από την γλώσσα υλοποίησης είτε λειτουργούν με την λογική των web services, προσφέροντας τα τεστ και το αποτέλεσμα του σε κάποια μορφή που είναι ανεξάρτητη από την γλώσσα υλοποίησης, συνήθως XML.

Κατά συνέπεια είχαμε αρκετά μεγάλη ελευθερία στην επιλογή των εργαλείων που θα χρησιμοποιηθούν. Προτιμήθηκε η γλώσσα PHP, καθώς πρόκειται για την πιο διαδεδομένη γλώσσα προγραμματισμού διαδικτυακών εφαρμογών.

5.2 Τεχνολογίες υλοποίησης

Παρακάτω παρουσιάζονται οι βασικές τεχνολογίες που χρησιμοποιήθηκαν για την κατασκευή της εφαρμογής.

5.2.1 HTML

Η HTML (HyperText Markup Language) είναι η κύρια γλώσσα περιγραφής για τις ιστοσελίδες, και τα στοιχεία της είναι τα βασικά δομικά στοιχεία των ιστοσελίδων. Ο σκοπός ενός web browser είναι να διαβάσει τα έγγραφα HTML και τα συνθέτει σε σελίδες που μπορεί κανείς να διαβάσει ή να ακούσει. Ο browser δεν εμφανίζει τις ετικέτες HTML, αλλά τις χρησιμοποιεί για να ερμηνεύσει το περιεχόμενο της σελίδας.

```
<!doctype HTML>
<html manifest="/cache.manifest">
  <head>
    <meta charset="utf-8" />
    <title>My Weblog</title>
    <link rel="stylesheet" href="style.css" />
  </head>

  <body>
    ...
  </body>
</html>
```

Εικόνα 5-1 Η βασική μορφή μίας σελίδας HTML

Η HTML γράφεται υπό μορφή στοιχείων HTML τα οποία αποτελούνται από ετικέτες(tags), οι οποίες περικλείονται μέσα σε σύμβολα «μεγαλύτερο από» και «μικρότερο από» (για παράδειγμα <html>), μέσα στο περιεχόμενο της ιστοσελίδας. Οι ετικέτες HTML συνήθως λειτουργούν ανά ζεύγη (για παράδειγμα <h1> και </h1>), με την πρώτη να ονομάζεται ετικέτα έναρξης και τη δεύτερη ετικέτα λήξης (ή σε άλλες περιπτώσεις ετικέτα ανοίγματος και ετικέτα κλεισίματος αντίστοιχα). Ανάμεσα στις ετικέτες, οι σχεδιαστές ιστοσελίδων μπορούν να τοποθετήσουν κείμενο, πίνακες, εικόνες κλπ.

Τα στοιχεία της HTML χρησιμοποιούνται για να κτίσουν όλους του ιστότοπους. Η HTML επιτρέπει την ενσωμάτωση εικόνων και άλλων αντικειμένων μέσα στη σελίδα, και μπορεί να χρησιμοποιηθεί για να εμφανίσει διαδραστικές φόρμες. Παρέχει τις μεθόδους δημιουργίας δομημένων εγγράφων (δηλαδή εγγράφων που αποτελούνται από το περιεχόμενο που μεταφέρουν και από τον κώδικα μορφοποίησης του περιεχομένου) καθορίζοντας δομικά σημαντικά στοιχεία για το κείμενο, όπως κεφαλίδες, παραγράφους, λίστες, συνδέσμους, παραθέσεις και άλλα. Μπορούν επίσης να ενσωματώνονται σενάρια εντολών σε γλώσσες όπως η JavaScript, τα οποία επηρεάζουν τη συμπεριφορά των ιστοσελίδων HTML.

Οι Web browsers μπορούν επίσης να αναφέρονται σε στυλ μορφοποίησης CSS (Cascading Style Sheets) για να ορίζουν την εμφάνιση και τη διάταξη του κειμένου και του υπόλοιπου υλικού. Ο οργανισμός W3C, ο οποίος δημιουργεί και συντηρεί τα πρότυπα για την HTML και τα CSS, ενθαρρύνει τη χρήση των CSS αντί διαφόρων στοιχείων της HTML για σκοπούς παρουσίασης του περιεχομένου.

5.2.2 JavaScript

Η JavaScript είναι μία γλώσσα προγραμματισμού ηλεκτρονικών υπολογιστών, η οποία έχει σαν σκοπό την παραγωγή δυναμικού περιεχομένου και την εκτέλεση κώδικα στην πλευρά του πελάτη (client-side) σε ιστοσελίδες. Το πρότυπο της γλώσσας κατά τον οργανισμό τυποποίησης ECMA ονομάζεται ECMAScript, ενώ δεν θα πρέπει να συγχέεται με τη Java, που είναι διαφορετική γλώσσα προγραμματισμού και με διαφορετικές εφαρμογές.

Η αρχική έκδοση της Javascript βασίστηκε στη σύνταξη στη γλώσσα προγραμματισμού C, αν και έχει εξελιχθεί, ενσωματώνοντας πια χαρακτηριστικά από νεότερες γλώσσες. Αρχικά χρησιμοποιήθηκε για προγραμματισμό από την πλευρά του πελάτη, δηλαδή στον browser του χρήστη, και χαρακτηρίζεται σαν client-side γλώσσα προγραμματισμού. Αυτό σημαίνει ότι η επεξεργασία του κώδικα Javascript και η παραγωγή του τελικού περιεχομένου HTML δεν πραγματοποιείται στο διακομιστή, αλλά στο πρόγραμμα περιήγησης των επισκεπτών, ενώ μπορεί να ενσωματωθεί σε στατικές σελίδες HTML. Αντίθετα, άλλες γλώσσες όπως η PHP εκτελούνται στο διακομιστή (server-side γλώσσες προγραμματισμού).

```
var displayClosure = function() {  
  var count = 0;  
  return function () {  
    return ++count;  
  };  
}  
var inc = displayClosure();  
inc(); // returns 1  
inc(); // returns 2  
inc(); // returns 3
```

Εικόνα 5-2 Η τυπική μορφή κώδικα JavaScript

Η JavaScript έχει γίνει μία από τις πιο δημοφιλείς γλώσσες προγραμματισμού ηλεκτρονικών υπολογιστών στον Παγκόσμιο Ιστό (Web). Με τη χρήση της τεχνολογίας Ajax, δηλαδή την δυναμική φόρτωση δεδομένων και τροποποίηση μίας ιστοσελίδας, η άλλαξε την παραδοσιακή δομή των ιστοσελίδων. Το αποτέλεσμα ήταν ένα καινοτόμο αντίκτυπο στην εξάπλωση των πλαισίων και των βιβλιοθηκών, τη βελτίωση προγραμματισμού με JavaScript, καθώς και αυξημένη χρήση της έξω από τα προγράμματα περιήγησης στο Web.

5.2.3 PHP

Η PHP είναι μια γλώσσα προγραμματισμού που σχεδιάστηκε για τη δημιουργία δυναμικών σελίδων στο διαδίκτυο και είναι επισήμως γνωστή ως: HyperText preprocessor. Είναι μια server-side (εκτελείτε στον διακομιστή) scripting γλώσσα που γράφεται συνήθως πλαισιωμένη από HTML, για μορφοποίηση των αποτελεσμάτων. Αντίθετα από μια συνηθισμένη HTML σελίδα η σελίδα PHP δεν στέλνεται άμεσα σε έναν πελάτη (client), αλλά πρώτα αναλύεται και μετά αποστέλλεται το παραγόμενο αποτέλεσμα. Τα στοιχεία HTML στον πηγαίο κώδικα μένουν ως έχουν, αλλά ο PHP κώδικας ερμηνεύεται και εκτελείται. Ο κώδικας PHP μπορεί να θέσει ερωτήματα σε βάσεις δεδομένων, να δημιουργήσει εικόνες, να διαβάσει και να γράψει αρχεία, να συνδεθεί με απομακρυσμένους υπολογιστές κα.

```
function getAdder($x) {  
    return function($y) use ($x) {  
        return $x + $y;  
    };  
}  
  
$adder = getAdder(8);  
echo $adder(2); // prints "10"
```

Εικόνα 5-3 Ένα παράδειγμα κώδικα PHP

Ο διακομιστής Apache, που χρησιμοποιείται σήμερα ευρέως σε συστήματα με τα λειτουργικά συστήματα GNU/Linux, Microsoft Windows, Mac OS X υποστηρίζει εξ ορισμού την εκτέλεση κώδικα PHP, είτε με την χρήση ενός πρόσθετου (mod_php) ή με την αποστολή του κώδικα προς εκτέλεση σε εξωτερική διεργασία CGI ή FCGI ή με την έλευση της php5.4 υποστηρίζονται η εκτέλεση σε πολυάσχολους ιστοχώρους, FastCGI Process Manager.

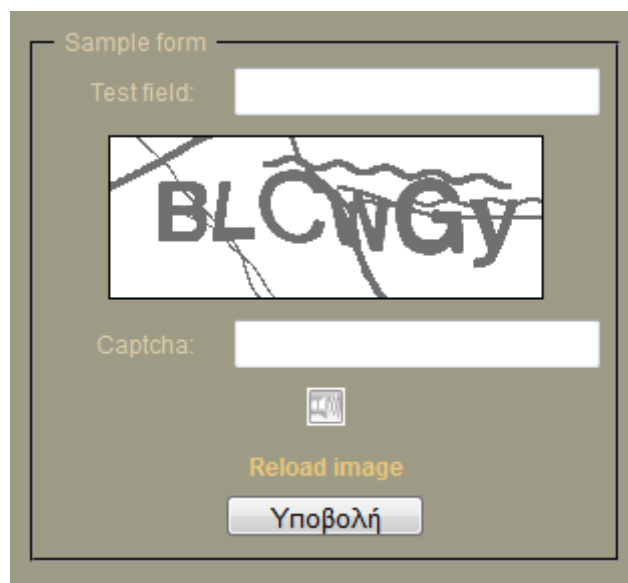
5.3 Δομή εφαρμογής

Η εφαρμογή που κατασκευάστηκε περιέχει ένα ευρετήριο με διαφορετικές εφαρμογές CAPTCHA που μπορούν να χρησιμοποιηθούν σε μία ιστοσελίδα. Σε κάθε μία από αυτές υπάρχει μία δοκιμαστική φόρμα με ένα πεδίο και το αντίστοιχο CAPTCHA. Παρακάτω θα παρουσιάσουμε τον τρόπο χρήσης κάθε ενός από αυτά για την ενσωμάτωσή του στη σελίδα και τον έλεγχο του αποτελέσματος.

5.3.1 Secureimage

5.3.1.1 Γενικά

Το Secureimage είναι μία από τις παλιότερες εφαρμογές CAPTCHA. Βασίζεται στην οπτική παραμόρφωση κάποιων γραμμάτων. Η δημιουργία του CAPTCHA γίνεται τοπικά στον server που το χρησιμοποιεί.



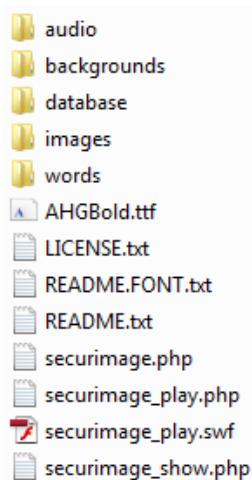
Εικόνα 5-4 Το CAPTCHA που δημιουργείται από το Secureimage μαζί με το κουμπί για το ακουστικό CAPTCHA και τον σύνδεσμο ανανέωσης της εικόνας.

Για να μπορέσει το Secureimage να δημιουργήσει τις εικόνες της δοκιμασίας απαιτεί ο διακομιστής να μπορεί να επεξεργαστεί αρχεία php. Η εγκατάσταση της γλώσσας php χρειάζεται να έχει γίνει με τέτοιο τρόπο ώστε να υποστηρίζεται η βιβλιοθήκη επεξεργασίας εικόνας GD.

Επιπλέον το Secureimage μπορεί να δημιουργήσει και captcha ήχου για τους χρήστες που έχουν δυσκολία στο διάβασμα της εικόνας. Αυτοί οι ήχοι αναπαράγονται στον υπολογιστή του χρήστη αν πατήσει το κουμπί με το megáφωνο μέσω μίας εφαρμογής Flash.

5.3.1.2 Εγκατάσταση

Η εφαρμογή έρχεται με αρκετά αρχεία που πρέπει να αντιγραφούν στον φάκελο της ιστοσελίδας.



Εικόνα 5-5 Τα αρχεία που απαιτεί το Secureimage για να λειτουργήσει.

Σε περίπτωση που είναι επιθυμητό ο προγραμματιστής μπορεί να αλλάξει αρκετές από τις παραμέτρους δημιουργίας του CAPTCHA μέσω του αρχείου securimage.php.

```
/**
 * The width of the captcha image
 * @var int
 */
public $image_width = 215;
/**
 * The height of the captcha image
 * @var int
 */
public $image_height = 80;
/**
 * The type of the image, default = png
 * @var int
 */
public $image_type = self::SI_IMAGE_PNG;

/**
 * The background color of the captcha
 * @var Securimage_Color
 */
public $image_bg_color = '#ffffff';
```

```

/**
 * The color of the captcha text
 * @var Securimage_Color
 */
public $text_color      = '#707070';
/**
 * The color of the lines over the captcha
 * @var Securimage_Color
 */
public $line_color      = '#707070';

```

Παραπάνω φαίνεται ένα απόσπασμα από αυτό το αρχείο. Αν είναι επιθυμητό μπορεί να αλλάξει το μέγεθος της εικόνας που δημιουργείται μέσω των μεταβλητών `image_width` και `image_height` και το είδος της εικόνας μέσω του `image_type`. Πέρα από αυτές τις επιλογές μπορούν να οριστούν τα χρώματα των γραμμάτων, του φόντου και των γραμμών, το πόσο παραμορφωμένοι και πολύπλοκοι θα είναι οι χαρακτήρες που θα παραχθούν, ποιοι χαρακτήρες θα μπορούν να περιέχονται στο τεστ και αρκετές ακόμη παράμετροι.

Έπειτα στην φόρμα που θα γίνει η εισαγωγή μπαίνει ένα στοιχείο `img`, το οποίο σαν πηγή έχει το αρχείο του `Secureimage`. Αξίζει να σημειωθεί ότι δεν παρέχεται σαν σύνδεσμος ένα τυπικό αρχείο εικόνας, αλλά ένα αρχείο `php`, το οποίο αναλαμβάνει να δημιουργήσει την εικόνα και να την στείλει. Επιπλέον είναι απαραίτητο να εισαχθεί και ένα πεδίο κειμένου στο οποίο ο χρήστης θα εισάγει την απάντησή του.

```

<form method="post">
  <fieldset>
    <legend>Sample form</legend>

    <label>Test field:</label><input type="text" />

    <center>
      <object type="application/x-shockwave-flash"
data="include/securimage/securimage_play.swf?audio_file=include/secur
image/securimage_play.php& bgColor1=fff& bgColor2=fff&
iconColor=777& borderWidth=1& borderColor=000" width="19"
height="19">
        <param name="movie"
value="include/securimage/securimage_play.swf?audio_file=/securimage/
securimage_play.php& bgColor1=fff& bgColor2=fff&
iconColor=777& borderWidth=1& borderColor=000" />
      </object>
    </center>

    <label>Captcha:</label><input type="text"
name="captcha_code" size="10" maxlength="6" />

    <a class="reloadImage" href="#"
onclick="document.getElementById('captcha').src =
'include/securimage/securimage_show.php?' + Math.random(); return

```



```

false">Reload image</a>

        <input type="submit" value="Υποβολή" />

    </fieldset>
</form>

```

Συμπληρωματικά υπάρχει έναν σύνδεσμος που δημιουργεί μία νέα εικόνα σε περίπτωση που ο χρήστης δεν μπορεί να καταλάβει την αρχική, με την χρήση javascript, χωρίς να χρειάζεται να ξαναφορτωθεί όλη η σελίδα. Επίσης υπάρχει η αναφορά στην εφαρμογή Flash που χρησιμοποιείται για το ακουστικό CAPTCHA.

Σε τέσσερα σημεία γίνεται αναφορά στα αρχεία του Secureimage, συγκεκριμένα στην ιδιότητα src του tag img, που περιγράφει την τοποθεσία της εικόνας, στη μεταβλητή src στον σύνδεσμο που ξαναφορτώνει την εικόνα και στα tag object και param που εμφανίζουν την εφαρμογή αναπαραγωγής ήχου. Αυτοί οι σύνδεσμοι χρειάστηκε να τροποποιηθούν για να δείχνουν στο φάκελο που έχουν τοποθετηθεί τα αρχεία της εφαρμογής, δηλαδή στην περίπτωση μας στο include/secureimage.

5.3.1.3 Επεξεργασία απάντησης

Όταν ο χρήστης πατήσει το κουμπί υποβολής της φόρμας χρειάζεται να ελεγχθεί αν η δοθείσα απάντηση είναι σωστή. Για να γίνει αυτό αρχικά χρειάζεται να φορτωθεί ένα απαραίτητο αρχείο και να δημιουργηθεί ένα αντικείμενο τύπου Secureimage.

```

require('include/secureimage/secureimage.php');
$secureimage = new Secureimage();

```

Έπειτα ελέγχεται αν το πεδίο captcha_code της φόρμας που περιέχει την απάντηση που έδωσε ο χρήστης περιέχει την σωστή απάντηση. Ο έλεγχος γίνεται χρησιμοποιώντας την συνάρτηση check του Secureimage, χωρίς να χρειάζεται επιπλέον κώδικας.

```

if ($secureimage->check($_REQUEST['captcha_code']) == false) {
    echo "<br /><h3>Ο κωδικός ασφαλείας που εισάγατε είναι λάθος..";
    echo "Παρακαλώ πηγαίνετε <a href='javascript:history.go(-1)'\>πίσω</a> και δοκιμάστε ξανά.</h3>";
} else {
    echo "<br /><h3>Η φόρμα υποβλήθηκε με επιτυχία.</h3>";
}

```

Ανάλογα με το αποτέλεσμα του ελέγχου της συνάρτησης check παρουσιάζεται στον χρήστη το κατάλληλο μήνυμα.

Ο πλήρης κώδικας της φόρμας δίνεται παρακάτω:

```

<?php include("include/header.php"); ?>

<h1>Secureimage</h1>

<p>Το Secureimage είναι μία από τις παλιότερες εφαρμογές CAPTCHA. Βασίζεται στην οπτική παραμόρφωση κάποιων γραμμάτων. Η δημιουργία του CAPTCHA γίνεται τοπικά στον server που το χρησιμοποιεί.</p>
<?php if(!isset($_REQUEST['captcha_code'])){ ?>
<form method="post">
  <fieldset>
    <legend>Sample form</legend>

    <label>Test field:</label><input type="text" />

    <center>
      <object type="application/x-shockwave-flash"
data="include/securimage/securimage_play.swf?audio_file=include/securimage/securimage_play.php&bgColor1=#fff&bgColor2=#fff&iconColor=#777&borderWidth=1&borderColor=#000" width="19"
height="19">
        <param name="movie"
value="include/securimage/securimage_play.swf?audio_file=/securimage/securimage_play.php&bgColor1=#fff&bgColor2=#fff&iconColor=#777&borderWidth=1&borderColor=#000" />
      </object>
    </center>

    <label>Captcha:</label><input type="text" name="captcha_code"
size="10" maxlength="6" />

    <a class="reloadImage" href="#"
onclick="document.getElementById('captcha').src =
'include/securimage/securimage_show.php?' + Math.random(); return
false">Reload image</a>

    <input type="submit" value="Υποβολή" />

  </fieldset>
</form>
<?php
}else{
  require('include/securimage/securimage.php');
  $securimage = new Securimage();

  if ($securimage->check($_REQUEST['captcha_code']) == false) {
    echo "<br /><h3>Ο κωδικός ασφαλείας που εισάγατε είναι
λάθος..";
    echo "Παρακαλώ πηγαίνετε <a href='javascript:history.go(-
1)'">πίσω</a> και δοκιμάστε ξανά.</h3>";
  }else{
    echo "<br /><h3>Η φόρμα υποβλήθηκε με επιτυχία.</h3>";
  }
}
}
?>

```

```
<?php include("include/footer.php"); ?>
```

5.3.1.4 Αξιολόγηση

Το Secureimage είναι μία εφαρμογή που τρέχει στον server της σελίδας που επιθυμεί να την χρησιμοποιήσει, δημιουργώντας κάθε φορά που ο χρήστης το επιθυμεί μία νέα εικόνα – τεστ. Αυτό έχει το πλεονέκτημα ότι μπορεί να χρησιμοποιηθεί σε σελίδες όπου δεν είναι επιθυμητή η χρήση κάποιας υπηρεσίας τρίτης εταιρείας για οποιονδήποτε λόγο, π.χ. για να μην υπάρχει πρόβλημα σε περίπτωση που η χρησιμοποιούμενη υπηρεσία δεν λειτουργεί για κάποιο λόγο.

Εξαιτίας αυτού η εγκατάσταση και η χρήση του έχει περισσότερα βήματα, δηλαδή πρέπει να αντιγραφούν τα αρχεία, να γίνουν οι ρυθμίσεις και να γραφτούν τα κατάλληλα links και includes στον κώδικα php και html αντίστοιχα, καθιστώντας την πολύπλοκη σε σχέση με άλλα αντίστοιχα εργαλεία. Καθώς τρέχει στον server της εταιρείας απαιτεί να είναι εγκαταστημένη η επέκταση GD της php η οποία αναλαμβάνει την δημιουργία και τροποποίηση εικόνων. Αν και συνήθως αυτό δεν είναι πρόβλημα καθώς πλέον αποτελεί μέρος της τυπικής εγκατάστασης της γλώσσας μπορεί να δημιουργήσει προβλήματα σε παλιότερα συστήματα που δεν είναι πρακτική η εγκατάσταση νέων επεκτάσεων.

Βέβαια η δυναμική δημιουργία μίας εικόνας CAPTCHA κάθε φορά που ζητείται μία νέα φόρμα απαιτεί αυξημένους υπολογιστικούς πόρους σε μνήμη και επεξεργαστική ισχύ. Αν και μία εικόνα δεν καταναλώνει αξιόλογους πόρους, σελίδες με μεγάλη κίνηση θα αντιμετωπίσουν μία υπολογίσιμη αύξηση του φόρτου το server τους. Αυτό μειώνει την χωρητικότητα του server, δηλαδή το πλήθος των αιτήσεων που μπορεί να εξυπηρετήσει ταυτόχρονα, και πιθανών να δημιουργήσει προβλήματα.

Ένα άλλο χαρακτηριστικό του που έχει ταυτόχρονα πλεονεκτήματα και μειονεκτήματα είναι η παρουσία του αρχείου ρυθμίσεων. Μέσω αυτού μπορεί να ρυθμιστεί ακριβώς το μέγεθος των εικόνων και το είδος τους, επιτρέποντας την πλήρη ενσωμάτωση της εικόνας στη σελίδα ώστε να παράγεται ένα αρτιότερο αισθητικό αποτέλεσμα. Επιπλέον μπορεί να καθοριστεί η δυσκολία που θα έχουν οι δοκιμασίες ώστε να μην είναι κουραστικές στους χρήστες. Όμως αυτό έχει το μειονέκτημα ότι, εκτός και αν ο προγραμματιστής που θα κάνει τις ρυθμίσεις έχει αρκετά εκτεταμένες γνώσεις, κινδυνεύει να δημιουργήσει εικόνες που θα είναι αρκετά εύκολο να διαβαστούν από προγράμματα OCR, χωρίς να είναι απαραίτητα εξίσου εύκολο να διαβαστούν από ανθρώπους.

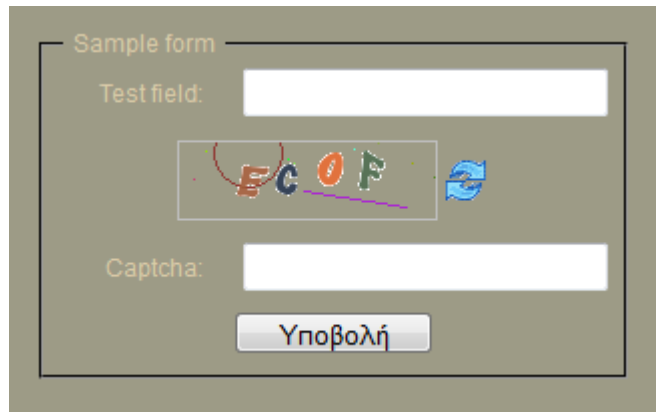
Πάντως το μεγάλο πλεονέκτημα του Secureimage είναι ότι παρέχει πέρα από την οπτική δοκιμασία αναγνώρισης των παραμορφωμένων γραμμάτων και μία ηχητική δοκιμασία για τους χρήστες που δεν επιθυμούν ή δεν μπορούν να επιλύσουν την πρώτη.

Ακόμη ένα πλεονέκτημα είναι πως αν και η βιβλιοθήκη είναι αρχικά ρυθμισμένη για να δίνει ηχητικές και οπτικές δοκιμασίες στα αγγλικά, είναι δυνατή η παραμετροποίηση και των δύο ώστε να μπορεί να χρησιμοποιηθεί εύκολα και από χρήστες άλλων εθνικοτήτων.

5.3.2 Cryptographp

5.3.2.1 Γενικά

Το Cryptographp είναι μία βιβλιοθήκη που παράγει γραφικά CAPTCHA που βασίζονται στην παραμόρφωση γραμμάτων. Οι εικόνες παράγονται τοπικά στον server που εγκαθίσταται.

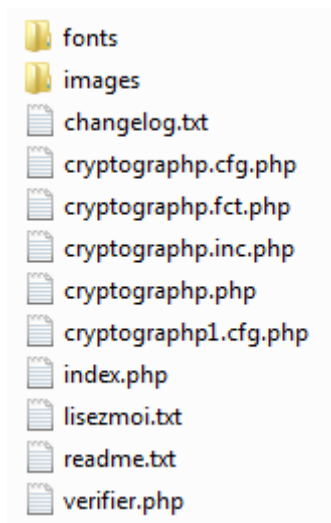


Εικόνα 5-6 Το captcha που παράγεται από την βιβλιοθήκη Cryptographp

Η βιβλιοθήκη είναι γραμμένη στην γλώσσα php, οπότε απαιτεί έναν server που μπορεί να εκτελέσει αρχεία php. Η επέκταση GD χρειάζεται να είναι εγκαταστημένη ώστε η εφαρμογή να μπορεί να επεξεργαστεί εικόνες.

5.3.2.2 Εγκατάσταση

Καθώς η βιβλιοθήκη εκτελείται τοπικά στον διακομιστή τα αρχεία της χρειάζεται να αντιγραφούν στον φάκελο της εφαρμογής.



Εικόνα 5-7 Τα αρχεία της εφαρμογής που χρειάζεται να αντιγραφούν

Στην περίπτωση της δικής μας εφαρμογής αντιγράφηκαν στον φάκελο include.

Το `cryptograph` δίνει την δυνατότητα πλήρους παραμετροποίησης μέσω του αρχείου ρυθμίσεων `cryptograph.cfg.php`. Μέσω αυτού είναι δυνατή η ρύθμιση της εικόνας που θα παραχθεί ώστε να ανταποκρίνεται στις απαιτήσεις του της σελίδας. Μερικές από τις ρυθμίσεις είναι το χρώμα των γραμμάτων και του φόντου, το ποια γράμματα θα περιέχονται στο `test` καθώς και η εφαρμογή πολύπλοκων εφέ στην εικόνα.

```
$charelc = 'BCDFGHKLMNPRTVWXZ'; // Consonnes utilisées si $crypteasy
= true
$charelv = 'AEIOUY'; // Voyelles utilisées si $crypteasy
= true

$difuplow = false; // Différencie les Maj/Min lors de la
saisie du code (true, false)

$charnbmin = 4; // Nb minimum de caracteres dans le
cryptogramme
$charnbmax = 4; // Nb maximum de caracteres dans le
cryptogramme

$charspace = 20; // Espace entre les caracteres (en pixels)
$charsizemin = 14; // Taille minimum des caractères
$charsizemax = 16; // Taille maximum des caractères

$charanglemax = 25; // Angle maximum de rotation des caracteres
(0-360)
$charup = true; // Déplacement vertical alternatif des caractères
(true/false)

// Effets supplémentaires

$cryptgaussianblur = false; // Transforme l'image finale en
brouillant: mode Gauss (true/false)
// uniquement si PHP >= 5.0.0
$cryptgrayscale = false; // Transforme l'image finale en
grayscale (true/false)
// uniquement si PHP >= 5.0.0
```

Παρόλο που οι ρυθμίσεις συνοδεύονται από αρκετά επεξηγηματικά σχόλια, αυτά είναι στα γαλλικά, κάτι αρκετά ασυνήθιστο καθώς οι βιβλιοθήκες ελεύθερου κώδικα παραδοσιακά είναι γραμμένες στα αγγλικά. Η επιλογή της γλώσσας προφανώς θα δυσκολέψει αρκετούς χρήστες που θα θελήσουν να παραμετροποιήσουν την βιβλιοθήκη.

Η ενσωμάτωση του `cryptographp` στην φόρμα γίνεται εισάγοντας αρχικά τα κατάλληλα αρχεία και ρυθμίσεις.

```
$cryptinstall = 'include/crypt/cryptographp.fct.php' ;  
include($cryptinstall);
```

Έπειτα χρειάζεται να εισαχθεί στην φόρμα η εικόνα και το κατάλληλο πεδίο. Η εικόνα εισάγεται με την χρήση της συνάρτησης `dsp_crypt(0,1)`. Η πρώτη παράμετρος αναφέρεται στο αρχείο ρυθμίσεων που θα χρησιμοποιηθεί. Αν είναι 0 φορτώνεται το αρχείο `cryptograph.cgf.php` που είναι το προκαθορισμένο αρχείο ρυθμίσεων, ενώ σε περίπτωση που είναι επιθυμητό να φορτωθεί ένα διαφορετικό αρχείο ρυθμίσεων τότε αυτό παρέχεται σαν αλφαριθμητικό. Η δεύτερη παράμετρος ρυθμίζει το αν θα εμφανίζεται το κουμπί επαναφόρτωσης της εικόνας, με το 1 να σημαίνει ότι εμφανίζεται, ενώ το 0 σημαίνει ότι δεν θα εμφανίζεται.

Πέρα από την εικόνα χρειάζεται να εισαχθεί και το πεδίο που θα κρατάει την απάντηση του χρήστη, όπως φαίνεται παρακάτω.

```
<form method="post">  
  <fieldset>  
    <legend>Sample form</legend>  
  
    <label>Test field:</label><input type="text" />  
  
    <?php dsp_crypt(0,1); ?>  
  
    <label>Captcha:</label><input type="text" name="code" size="10"  
maxlength="6" />  
  
    <input type="submit" value="Υποβολή" />  
  
  </fieldset>  
</form>
```

5.3.2.3 Επεξεργασία απάντησης

Για να επεξεργαστεί η απάντηση που έδωσε ο χρήστης είναι απαραίτητη η εισαγωγή των κατάλληλων αρχείων.

```
$cryptinstall = 'include/crypt/cryptographp.fct.php' ;  
include($cryptinstall);
```

Ο έλεγχος γίνεται περνώντας στην συνάρτηση `chk_crypt` σαν παράμετρο την απάντηση που έδωσε ο χρήστης.

```

    if (chk_crypt($_REQUEST['code']) == false) {
        echo "<br /><h3>Ο κωδικός ασφαλείας που εισάγατε είναι
λάθος..";
        echo "Παρακαλώ πηγαίνετε <a href='javascript:history.go(-
1)''>πίσω</a> και δοκιμάστε ξανά.</h3>";
    }else{
        echo "<br /><h3>Η φόρμα υποβλήθηκε με επιτυχία.</h3>";
    }
}

```

Αυτή επιστρέφει true ή false ανάλογα με το αν έχει δοθεί σωστή η λάθος απάντηση, οπότε και εμφανίζεται το κατάλληλο μήνυμα.

Ο πλήρης κώδικας της φόρμας φαίνεται παρακάτω:

```

<?php
include("include/header.php");
$cryptinstall = 'include/crypt/cryptographp.fct.php';
include($cryptinstall);
?>

<h1>Cryptographp</h1>

<p>Το Cryptographp είναι μία βιβλιοθήκη που παράγει γραφικά CAPTCHA
που βασίζονται στην παραμόρφωση γραμμάτων.</p>
<?php if(!isset($_REQUEST['code'])) { ?>
<form method="post">
    <fieldset>
        <legend>Sample form</legend>

        <label>Test field:</label><input type="text" />

        <?php dsp_crypt(0,1); ?>

        <label>Captcha:</label><input type="text" name="code" size="10"
maxlength="6" />

        <input type="submit" value="Υποβολή" />

    </fieldset>
</form>
<?php
}else{

    if (chk_crypt($_REQUEST['code']) == false) {
        echo "<br /><h3>Ο κωδικός ασφαλείας που εισάγατε είναι
λάθος..";
        echo "Παρακαλώ πηγαίνετε <a href='javascript:history.go(-
1)''>πίσω</a> και δοκιμάστε ξανά.</h3>";
    }else{
        echo "<br /><h3>Η φόρμα υποβλήθηκε με επιτυχία.</h3>";
    }
}
}
?>

```

```
<?php include("include/footer.php"); ?>
```

5.3.2.4 Αξιολόγηση

Η βιβλιοθήκη `cryptographr` έχει αρκετές ομοιότητες με το `Secureimage` που μελετήσαμε νωρίτερα. Συγκεκριμένα πρόκειται για μία βιβλιοθήκη γραμμένη σε `php` που τρέχει τοπικά στον διακομιστή του της σελίδας που θα το χρησιμοποιήσει, απαιτώντας συγκεκριμένες ρυθμίσεις από αυτόν και ανεβάζοντας τον φόρτο εργασίας του.

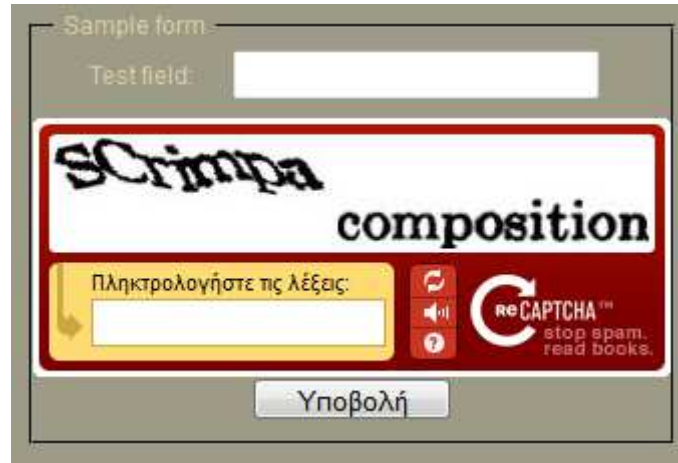
Το αρχείο ρυθμίσεων δίνει την δυνατότητα παραμετροποίησης των εικόνων που παράγονται. Οι ρυθμίσεις είναι εκτενέστερες και κάποιες από αυτές είναι αρκετά προηγμένες, όπως για παράδειγμα η εφαρμογή φίλτρου `Gaussian Blur` ή η ρύθμιση του θορύβου που θα περιέχει η παραγόμενη εικόνα. Αν και πρόκειται για εντυπωσιακές επιλογές, αξίζει να σημειωθεί ότι έχουν και σοβαρή επίπτωση στην επεξεργαστική ισχύ που καταναλώνεται για την παραγωγή των εικόνων. Επίσης η τροποποίηση των ρυθμίσεων από κάποιον που δεν έχει γνώση των βασικών αρχών της επεξεργασίας εικόνας και της οπτικής ανάλυσης χαρακτήρων μπορεί να δημιουργήσει στην δημιουργία CAPTCHA που είναι πολύ απλά να επιλυθούν από τους κατάλληλους αλγορίθμους, ενώ ταυτόχρονα μπορεί να είναι αναίτια δύσκολα για τους χρήστες.

5.3.3 reCAPTCHA

5.3.3.1 Γενικά

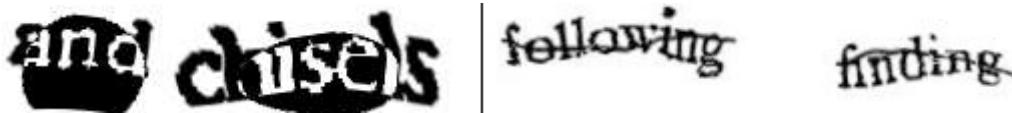
Το `reCAPTCHA` είναι ένα τεστ παραμόρφωσης χαρακτήρων το οποίο αναπτύχθηκε στο `Carnegie Mellon University` και έπειτα αγοράστηκε από την εταιρεία `Google`. Διαφέρει σε σχέση με τις προηγούμενες βιβλιοθήκες που εξετάσαμε στο ότι δεν τρέχει τοπικά στον `server` που το χρησιμοποιεί αλλά παρέχεται σαν υπηρεσία. Πρόκειται για το πιο δημοφιλές τεστ CAPTCHA σήμερα, παρουσιάζοντας εκατοντάδες εκατομμύρια τεστ καθημερινά.

Οι λέξεις που παρουσιάζονται στους χρήστες δεν παράγονται τυχαία, αλλά παράγονται μέσω της διαδικασίας ψηφιοποίησης κανονικών βιβλίων, σκοπός της οποίας είναι η μετατροπή κανονικών βιβλίων σε ηλεκτρονική μορφή. Αυτό γίνεται παίρνοντας φωτογραφίες κάθε σελίδας του βιβλίου και περνώντας τις σαν είσοδο σε δύο ανεξάρτητα προγράμματα οπτικής αναγνώρισης χαρακτήρων. Τα δύο διαφορετικά αποτελέσματα (κείμενα) που παράγουν τα προγράμματα για κάθε σελίδα συγκρίνονται μεταξύ τους και επιλέγονται όσες λέξεις διαφέρουν μεταξύ τους και όσες δεν υπάρχουν στο λεξικό. Αυτές χρησιμοποιούνται στο τεστ CAPTCHA που παράγεται.



Εικόνα 5-8 Το τεστ reCAPTCHA όπως εμφανίζεται σε μία φόρμα.

Συγκεκριμένα κάθε δοκιμασία παρουσιάζει στον χρήστη δύο εικόνες που απεικονίζουν δύο λέξεις, μία από τις οποίες είναι γνωστή και μία άγνωστη. Αξίζει να σημειωθεί ότι αυτές οι εικόνες δεν είναι αυτούσιες όπως παρουσιάζονται στο βιβλίο αλλά μπορεί να έχουν επιπλέον παραμόρφωση. Όταν ο χρήστης απαντήσει στη δοκιμασία συγκρίνεται η απάντηση που έχει δώσει στην γνωστή λέξη και αν είναι σωστή αποθηκεύεται η απάντησή που έχει δώσει στην άγνωστη (η οποία δεν λαμβάνεται υπόψη στο αν έχει απαντήσει σωστά στο τεστ), και όταν μία συγκεκριμένη απάντηση συγκεντρώσει ένα συγκεκριμένο πλήθος τότε θεωρείται σωστή και στο μέλλον χρησιμοποιείται σαν γνωστή λέξη.



Εικόνα 5-9 Διαφορετικές τεχνικές παραμόρφωσης που έχουν κατά καιρούς χρησιμοποιηθεί στα τεστ του reCAPTCHA.

Το γεγονός πως το reCAPTCHA είναι τόσο δημοφιλές το έχει κάνει κατά καιρούς στόχο πολλών επιθέσεων, και έχει οδηγήσει στην αλλαγή του τρόπου παραμόρφωσης των γραμμάτων του τεστ. Όμως, εξαιτίας της φύσης του τεστ το μεγαλύτερο ποσοστό επιτυχίας που είχε επιτευχθεί ήταν 31,8%, και μέσα σε διάστημα λίγων ορών η εταιρεία άλλαξε το τεστ με τέτοιο τρόπο ώστε να ξεπεραστεί το πρόβλημα.

Το reCAPTCHA περιέχει και μία ηχητική δοκιμασία για τους χρήστες με προβλήματα όρασης η οποία στο παρελθόν έχει χρησιμοποιηθεί από επιτιθέμενους για να σπάσει το τεστ, καθώς ήταν πολύ ευκολότερο από το οπτικό τεστ. Την άνοιξη του 2012 μία ομάδα προγραμματιστών έκανε επίδειξη μίας τεχνικής η οποία με την χρήση ενός αλγορίθμου τεχνητής νοημοσύνης ανέλυε το ηχητικό μήνυμα και είχε ποσοστό επιτυχίας 99,1%. Το Google άλλαξε άμεσα το τεστ με την ομάδα να αλλάζει και τον δικό της αλγόριθμο ώστε να έχει ποσοστό επιτυχίας γύρω στο 60%.

Μετά από δύο ακόμη τέτοιες αλλαγές το Google έχει κάνει αρκετά δυσκολότερο το ηχητικό CAPTCHA, με την διάρκεια του τεστ να έχει αυξηθεί στα 30 δευτερόλεπτα από 8 που ήταν αρχικά και την παραμόρφωση κάθε ήχου να είναι εξαιρετικά έντονη. Αυτή τη

στιγμή τα ηχητικά CAPTCHA του reCAPTCHA δεν είναι δυνατόν να επιλυθούν από αλγόριθμους αναγνώρισης φωνής αλλά είναι επίσης σχεδόν αδύνατον να επιλυθούν από ανθρώπους, καθιστώντας τα πρακτικά άχρηστα.

5.3.3.2 Εγκατάσταση

5.3.3.2.1 Δημιουργία κλειδιών

Καθώς το reCAPTCHA παρέχεται σαν υπηρεσία η διαδικασία εγκατάστασής του είναι αρκετά διαφορετική σε σχέση με τις λύσεις που εξετάσαμε προηγουμένως.

Πριν την ίδια την εγκατάσταση είναι απαραίτητη η δημιουργία ενός λογαριασμού στο google.com. Ο λόγος είναι πως για να επιτρέπεται η χρήση του reCAPTCHA χρειάζεται η έκδοση ενός ειδικού κλειδιού από το Google. Έπειτα είναι απαραίτητη η δημιουργία αυτού του κλειδιού για τη σελίδα που θα το χρησιμοποιεί, η οποία γίνεται από την διεύθυνση <https://www.google.com/recaptcha/admin/create>.



Εικόνα 5-10 Η σελίδα δημιουργίας νέου κλειδιού για το reCAPTCHA.

Σε αυτή την φόρμα εισάγαμε την διεύθυνση localhost, καθώς η εφαρμογή θα τρέχει τοπικά στον υπολογιστή. Αν πρόκειται για μία σελίδα που έχει δικό της URL τότε είναι απαραίτητη η εισαγωγή αυτής της διεύθυνσης στην φόρμα. Επιπλέον μέσω της επιλογής Enable this key on all domains (global key) μπορεί να δημιουργηθεί ένα κλειδί που θα λειτουργεί σε όλες τις σελίδες του χρήστη, άσχετα από την διεύθυνσή τους.



Εικόνα 5-11 Η λίστα με τους ιστοτόπους για τους οποίους έχουν δημιουργηθεί κλειδιά για χρήση μέσω του reCAPTCHA. Εδώ φαίνεται η διεύθυνση localhost, την οποία μόλις δημιουργήσαμε.

Μετά την δημιουργία του κλειδιού φαίνεται μία λίστα με όλους τους ιστοτόπους που έχουν δημιουργηθεί, δηλαδή σε αυτή την περίπτωση μας τον localhost. Επιλέγοντας τον εμφανίζονται τα κλειδιά που έχουν συσχετιστεί με αυτόν και πληροφορίες σχετικά με την χρήση τους.



Εικόνα 5-12 Τα κλειδιά που έχουν συσχετιστεί με τον ιστότοπο localhost.

Τα παραπάνω κλειδιά θα χρησιμοποιηθούν για την επικοινωνία του server με το Google. Καθώς η υπηρεσία δεν τρέχει στον server της σελίδας είναι απαραίτητη η ασφαλής επικοινωνία του με το Google, χωρίς να μπορεί να παρεμβληθεί ένας τρίτος και να αλλοιώσει τα αποτελέσματα. Οι απαντήσεις στέλνονται κωδικοποιημένες και τα δύο κλειδιά που χρησιμοποιούνται είναι το ιδιωτικό και το δημόσιο κλειδί, ώστε να μπορεί να διασφαλιστεί η ασφαλής επικοινωνία μέσω της ασύμμετρης κρυπτογράφησης των δεδομένων που στέλνονται.

Τα διαφορετικά κλειδιά εξυπηρετούν στην διασφάλιση της ασφάλειας της επικοινωνίας σε περίπτωση που τα κλειδιά μίας συγκεκριμένης σελίδας πέσουν στα χέρια του επιτιθέμενου. Με αυτόν τον τρόπο διασφαλίζεται πως οι υπόλοιπες σελίδες του ίδιου χρήστη δεν θα είναι ευάλωτες σε επίθεση.

5.3.3.2.2 Εγκατάσταση αρχείων και κώδικας

Καθώς το reCAPTCHA δεν τρέχει τοπικά στον server που το χρησιμοποιεί, δεν χρειάζεται να αντιγραφούν σε αυτόν τα αρχεία της εφαρμογής. Όμως για να γίνει η επικοινωνία ανάμεσα στους δύο servers χρειάζεται μία αρκετά πολύπλοκη διαδικασία ώστε να εμφανιστεί το CAPTCHA στην σελίδα και να επιβεβαιωθεί η απάντηση, χρησιμοποιώντας ασφαλείς συνδέσεις για τη συνομιλία.

Για την διευκόλυνση της επικοινωνίας διατίθεται ένα μεγάλο πλήθος βιβλιοθηκών που αυτοματοποιούν την διαδικασία. Υπάρχουν επιλογές για πάρα πολλές γλώσσες όπως php, Java και JSP, ASP και ASP.NET και PERL. Επιπλέον διατίθενται έτοιμα πρόσθετα για αρκετές έτοιμες εφαρμογές, όπως WordPress, MediaWiki, phpBB κα. Για τις γλώσσες που δεν υποστηρίζονται υπάρχουν λεπτομερείς οδηγίες σχετικά με το πώς πρέπει να γίνεται η επικοινωνία.

Στην περίπτωσή μας επιλέχθηκε η χρήση της έτοιμης βιβλιοθήκης για την γλώσσα PHP, δηλαδή του αρχείου recaptchalib.php, το οποίο αντιγράφηκε στον φάκελο includes της εφαρμογής.

Για την εμφάνιση του CAPTCHA χρειάζεται να μπει ο κατάλληλος κώδικας PHP κατά την δημιουργία της σελίδας. Αυτός συμπεριλαμβάνει το αρχείο recaptchalib.php, για το οποίο χρειάζεται να προσδιορίσουμε ότι βρίσκεται στον φάκελο include. Έπειτα ορίζεται μια μεταβλητή που περιέχει το δημόσιο κλειδί της εφαρμογής και καλείται η συνάρτηση recaptcha_get_html με παράμετρο την παραπάνω μεταβλητή, η οποία επιστρέφει το CAPTCHA, το οποίο έπειτα εμφανίζεται.

```
<form method="post">
  <fieldset>
    <legend>Sample form</legend>

    <label>Test field:</label><input type="text" />

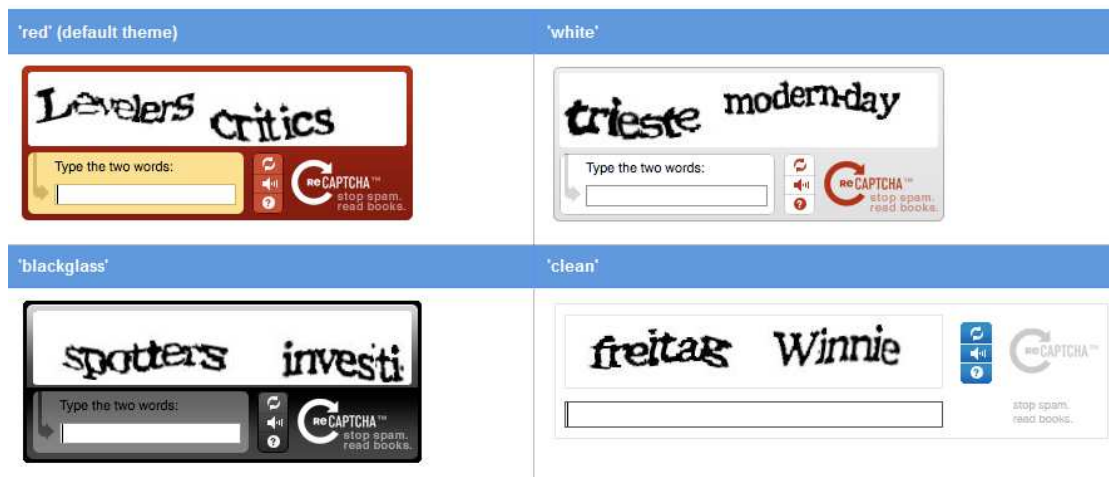
    <?php
      require_once('include/recaptcha/recaptchalib.php');
      $publickey = "6LdTJNASAAAAAKOorSRdzNbdKOWRJVXcyC7RKMA1";
      echo recaptcha_get_html($publickey);
    ?>

    <input type="submit" value="Υποβολή" />

  </fieldset>
</form>
```

Σε αντίθεση με άλλες μεθόδους το reCAPTCHA δεν χρειάζεται την εισαγωγή επιπλέον πεδίου για να εισάγει ο χρήστης την απάντηση, καθώς αυτό ήδη περιέχει το απαιτούμενο πεδίο.

Καθώς το reCAPTCHA εμφανίζει όλο το πλαίσιο διαλόγου, αυτό μπορεί να μην ταιριάζει αισθητικά με την εμφάνιση της υπόλοιπης σελίδας. Για να ξεπεραστεί αυτό το εμπόδιο ο χρήστης μπορεί να χρησιμοποιήσει κάποιο από τα έτοιμα θέματα που φαίνονται παρακάτω.



Εικόνα 5-13 Τα έτοιμα θέματα που υποστηρίζει το reCAPTCHA

Για την χρήση ενός από αυτά, πέρα από το κόκκινο που είναι το προεπιλεγμένο είναι απαραίτητο να εισαχθεί επιπλέον ένα απόσπασμα κώδικα JavaScript στην σελίδα.

```
<script type="text/javascript">
  var RecaptchaOptions = {
    theme : 'white'
  };
</script>
```

Στον πίνακα που δημιουργείται στην επιλογή theme εισάγεται το κατάλληλο θέμα.

Αν τα παραπάνω θέματα δεν είναι αρκετά τότε υπάρχουν αρκετές περισσότερες επιλογές με τις οποίες είναι δυνατόν να τροποποιηθεί εντελώς η εμφάνιση του, όμως η ανάλυσή τους ξεφεύγει από τα πλαίσια της εργασίας.

Πέρα από την εμφάνιση είναι δυνατόν να ρυθμιστούν και τα μηνύματα που εμφανίζει το reCAPTCHA. Υπάρχουν έτοιμες μεταφράσεις για κάποιες γλώσσες, ενώ επιπλέον παρέχονται οδηγίες για την εισαγωγή μηνυμάτων σε οποιαδήποτε γλώσσα.

Which language is used in the interface for the pre-defined themes. The following languages are supported:

Language	Code
English	en
Dutch	nl
French	fr
German	de
Portuguese	pt
Russian	ru
Spanish	es
Turkish	tr

Εικόνα 5-14 Οι γλώσσες που υποστηρίζονται από το reCAPTCHA.

Ο πλήρης κώδικας της φόρμας φαίνεται παρακάτω:

```
<?php include("include/header.php");
// private: 6LdTJNASAAAAAABO5k0hBSr_ZSP5HSmKyoTXjZJjR
// public : 6LdTJNASAAAAAKOorSRdzNbdKOWRJVXcyC7RKMA1
?>

<h1>reCAPTCHA</h1>

<p>Το reCAPTCHA είναι ένα τεστ παραμόρφωσης χαρακτήρων το οποίο αναπτύχθηκε στο Carnegie Mellon University και έπειτα αγοράστηκε από την εταιρεία Google. Διαφέρει σε σχέση με τις προηγούμενες βιβλιοθήκες που εξετάσαμε στο ότι δεν τρέχει τοπικά στον server που το χρησιμοποιεί αλλά παρέχεται σαν υπηρεσία. Πρόκειται για το πιο δημοφιλές τεστ CAPTCHA σήμερα, παρουσιάζοντας εκατοντάδες εκατομμύρια τεστ καθημερινά.</p>

<?php if(!isset($_REQUEST['recaptcha_response_field'])) { ?>
<form method="post">
<fieldset>
```

```

<legend>Sample form</legend>

<label>Test field:</label><input type="text" />

<?php
    require_once('include/recaptcha/recaptchalib.php');
    $publickey = "6LdTJNASAAAAAKOorSRdzNbdKOWRJVXcyC7RKMA1";
    echo recaptcha_get_html($publickey);
?>

<input type="submit" value="Υποβολή" />

</fieldset>
</form>
<?php
}else{
    require_once('include/recaptcha/recaptchalib.php');
    $privatekey = "6LdTJNASAAAAABO5k0hBSr_ZSP5HSmKyoTXjZJjR";
    $resp = recaptcha_check_answer ($privatekey,
        $_SERVER["REMOTE_ADDR"],
        $_REQUEST["recaptcha_challenge_field"],
        $_REQUEST["recaptcha_response_field"]);

    if (!$resp->is_valid) {
        echo "<br /><h3>Ο κωδικός ασφαλείας που εισάγατε είναι
λάθος..";
        echo "Παρακαλώ πηγαίνετε <a href='javascript:history.go(-
1)'\>πίσω</a> και δοκιμάστε ξανά.</h3>";
    }else{
        echo "<br /><h3>Η φόρμα υποβλήθηκε με επιτυχία.</h3>";
    }
}
?>

<?php include("include/footer.php"); ?>

```

5.3.3.3 Επεξεργασία απάντησης

Για την επεξεργασία της απάντησης χρησιμοποιείται το ιδιωτικό κλειδί που έχει παραχθεί κατά τη δήλωση της σελίδας στο Google. Χρειάζεται να περαστεί σαν όρισμα στην συνάρτηση `recaptcha_check_answer` μία μεταβλητή που να περιέχει το ιδιωτικό κλειδί, καθώς και κάποια από τα πεδία της φόρμας, τα οποία έχουν δημιουργηθεί αυτόματα κατά την εισαγωγή της δοκιμασίας της φόρμας.

```

require_once('include/recaptcha/recaptchalib.php');
$privatekey = "6LdTJNASAAAAABO5k0hBSr_ZSP5HSmKyoTXjZJjR";
$resp = recaptcha_check_answer ($privatekey,
    $_SERVER["REMOTE_ADDR"],
    $_REQUEST["recaptcha_challenge_field"],
    $_REQUEST["recaptcha_response_field"]);

```

Η συνάρτηση επιστρέφει την μεταβλητή `resp`, η οποία μεταξύ άλλων περιέχει και την μεταβλητή `is_valid`. Η τιμή της είναι `true` ή `false`, ανάλογα με το αν δόθηκε σωστή ή λάθος απάντηση

```
if (!$resp->is_valid) {  
    echo "<br /><h3>Ο κωδικός ασφαλείας που εισάγατε είναι  
λάθος..";  
    echo "Παρακαλώ πηγαίνετε <a href='javascript:history.go(-  
1)''>πίσω</a> και δοκιμάστε ξανά.</h3>";  
}else{  
    echo "<br /><h3>Η φόρμα υποβλήθηκε με επιτυχία.</h3>";  
}
```

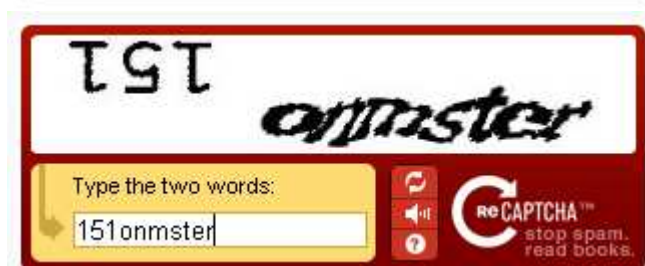
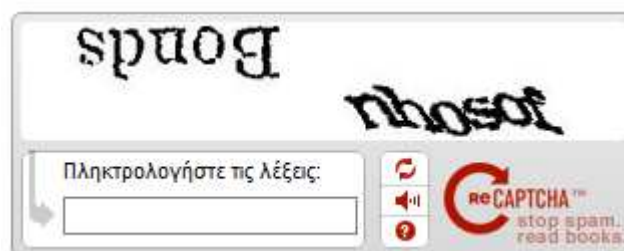
Η απάντηση ελέγχεται και ανάλογα με το αν είναι σωστή ή λάθος εμφανίζεται το κατάλληλο μήνυμα.

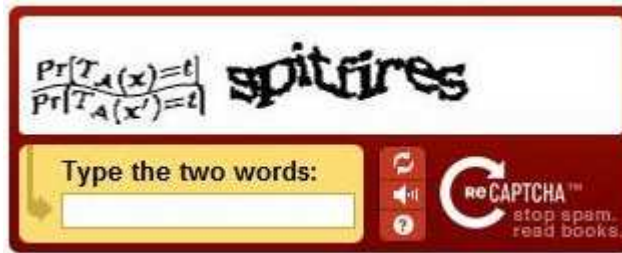
5.3.3.4 Αξιολόγηση

Το reCAPTCHA διαφέρει πάρα πολύ σε σχέση με τις υπόλοιπες βιβλιοθήκες που αναλύσαμε μέχρι τώρα από τεχνολογική άποψη. Όμως η μεγαλύτερη διαφορά πηγάζει από το ότι το αναπτύσσει και το συντηρεί η εταιρεία Google. Αυτό φαίνεται και από την ποιότητα της παρεχόμενης υπηρεσίας και από το πλήθος των επιλογών που παρέχονται.

Σε καθαρά τεχνολογικό επίπεδο το μεγαλύτερο πλεονέκτημα που έχει η συγκεκριμένη βιβλιοθήκη είναι πως χρησιμοποιεί λέξεις οι οποίες έχουν ήδη αποτύχει να αναγνωριστούν από προγράμματα αναγνώρισης χαρακτήρων. Αυτές οι λέξεις παραμορφώνονται σε ακόμη μεγαλύτερο βαθμό κάνοντας εξαιρετικά δύσκολη την αναγνώρισή τους.

Βέβαια μία παρενέργεια αυτού του τρόπου δημιουργίας των τεστ είναι πως μπορεί να δοθούν κάποιες ερωτήσεις οι οποίες δεν είναι δυνατών να απαντηθούν, όπως αυτές που φαίνονται παρακάτω.





Εικόνα 5-15 Κάποια παραδείγματα τεστ που παράγονται από το reCAPTCHA τα οποία δεν είναι δυνατόν να επιλυθούν.

Παρόλα αυτά καθώς υπάρχουν εκατομμύρια πιθανές ερωτήσεις οι συγκεκριμένες ερωτήσεις εμφανίζονται πολύ σπάνια.

Το reCAPTCHA παρέχει επιπλέον και ένα ηχητικό τεστ αλλά όπως αναλύσαμε προηγουμένως καθώς έχει γίνει επανειλημμένα στόχος επιθέσεων πλέον έχει γίνει τόσο δύσκολο ώστε να μην μπορεί να απαντηθεί ούτε από ανθρώπους.

Γενικότερα το γεγονός πως είναι τόσο δημοφιλές λειτουργεί ταυτόχρονα σαν πλεονέκτημα και μειονέκτημα γιατί αφού χρησιμοποιείται σε τόσες πολλές σελίδες γίνεται συνεχώς στόχος επιθέσεων. Αυτό οδηγεί στην συνεχή βελτίωσή του ώστε να αντιστέκεται στις επιθέσεις, αλλά όπως είδαμε προηγουμένως αυτό έχει οδηγήσει στην αχρήστευση του ηχητικού τμήματός του.

Σε καθαρά τεχνικό επίπεδο το ότι παρέχεται σαν υπηρεσία δίνει την δυνατότητα χρήσης εκατομμυρίων διαφορετικών δοκιμασιών χωρίς να επιβαρύνεται καθόλου ο server που το χρησιμοποιεί, καθώς δεν χρειάζεται να αποθηκεύονται σε αυτόν ούτε χρειάζεται να καταναλώνεται επεξεργαστική ισχύς για την δημιουργία του.

Το γεγονός πως για να φορτωθεί η δοκιμασία και να επιβεβαιωθεί η εγκυρότητα της απάντησης χρειάζεται να επικοινωνήσει με τους διακομιστές όπου φιλοξενείται το reCAPTCHA δημιουργεί το πιθανό πρόβλημα πως σε περίπτωση που υπάρχει κάποιο πρόβλημα με αυτούς η υπηρεσία δεν θα είναι διαθέσιμη. Όμως, καθώς η υποδομή που διαφέρει το Google είναι μία από τις μεγαλύτερες του διαδικτύου είναι γενικά απίθανο να συμβεί κάτι τέτοιο ή, σε περίπτωση που όντως υπάρξει διακοπή της υπηρεσίας να κρατήσει για αρκετά μεγάλο χρονικό διάστημα.

Οι απαιτήσεις που πρέπει να πληρεί ο server είναι ελάχιστες ενώ παρέχονται έτοιμες βιβλιοθήκες με τη χρήση των οποίων είναι δυνατή η εύκολη ενσωμάτωση του στην πλειοψηφία σε οποιαδήποτε εφαρμογή. Η μοναδική απαίτηση που υπάρχει από τον χρήστη είναι να έχει ενεργοποιημένη την εκτέλεση της γλώσσας JavaScript καθώς χρησιμοποιείται από την βιβλιοθήκη για την ανανέωση της εικόνας και την αναπαραγωγή του ηχητικού τεστ. Αξίζει να σημειωθεί πως για την αναπαραγωγή ήχου δεν χρησιμοποιείται η τεχνολογία Flash.

Στον τομέα της πολυγλωσσικής υποστήριξης η βιβλιοθήκη αποτυγχάνει να υποστηρίξει οποιαδήποτε άλλη γλώσσα εκτός από τα αγγλικά, καθώς προβάλλει λέξεις από αγγλικά βιβλία και δεν υπάρχει επιλογή για την χρήση οποιασδήποτε άλλης γλώσσας, όπως και για το ηχητικό τεστ. Όμως το κείμενο που εμφανίζεται για να δώσει τις οδηγίες στον

χρήστη μπορεί να αλλάξει και υπάρχουν έτοιμες μεταφράσεις για μία μεγάλη πληθώρα γλωσσών.

Τέλος το γεγονός πως οι απαντήσεις που δίνονται δεν είναι τυχαίες φράσεις αλλά παράγουν ένα χρήσιμο αποτέλεσμα, βοηθώντας στην ψηφιοποίηση βιβλίων, μπορεί να κάνει τους χρήστες να ξεπεράσουν την ενόχληση που μπορεί να νοιώσουν επιλύοντας το. Όμως κατά την επίλυση του δεν είναι προφανής ο τρόπος με τον οποίο δημιουργούνται τα τεστ, κάτι που θα μπορούσε να διορθωθεί.

5.3.4 Asirra

5.3.4.1 Γενικά

Το Asirra (Animal Species Image Recognition for Restricting Access) είναι ένα captcha αναγνώρισης εικόνων που είναι προϊόν έρευνας των εργαστηρίων της εταιρείας Microsoft. Η βασική λειτουργία του είναι ότι προβάλλει στον χρήστη μία σειρά από εικόνες αδέσποτων ζώων, τις οποίες παίρνει από τη σελίδα Petfinder και τον προτρέπει να επιλέξει όλα τα ζώα ίδιου τύπου.



Εικόνα 5-16 Η δοκιμασία του CAPTCHA Asirra.

Η ενσωμάτωση της δοκιμασίας γίνεται με χρήση JavaScript και η επαλήθευση γίνεται μέσω ενός Web Service.

5.3.4.2 Εγκατάσταση

Για την ενσωμάτωση του Asirra χρειάζεται να συμπεριληφθεί το κατάλληλο αρχείο JavaScript.

```
<script src="http://challenge.asirra.com/js/AsirraClientSide.js">
</script>
```

Έπειτα στο σημείο που πρέπει να εμφανιστεί το captcha πρέπει να καλεσθεί μία μέθοδος που καθορίζει το πλήθος των κελιών ανά γραμμή του πίνακα.

```
<script>
asirraState.SetCellsPerRow(4);
</script>
```

Στη συνέχεια πρέπει να γραφτεί μία μέθοδος που θα ελέγχει αν ο χρήστης έδωσε σωστά αποτελέσματα.

```
function MySubmitForm() {
  if (passThroughFormSubmit) {
    return true;
  }
  // Do site-specific form validation here, then...
  Asirra_CheckIfHuman(HumanCheckComplete);
  return false;
}

function HumanCheckComplete(isHuman) {
  if (!isHuman) {
    alert("Please correctly identify the cats.");
  }
  else {
    passThroughFormSubmit = true;
    formElt = document.getElementById("mainForm");
    formElt.submit();
  }
}
</script>
```

Αυτή η μέθοδος πρέπει να καλείται όταν η φόρμα πάει να γίνει submit.

```
<form method="post" onsubmit="return MySubmitForm();" id="mainForm">
```

Άρα ο συνολικό κώδικας της σελίδας που χρειάζεται να περιλαμβάνει όλα τα παραπάνω είναι ο εξής:

```

<script type="text/javascript">
var passThroughFormSubmit = false;
function MySubmitForm() {
  if (passThroughFormSubmit) {
    return true;
  }
  // Do site-specific form validation here, then...
  Asirra_CheckIfHuman(HumanCheckComplete);
  return false;
}

function HumanCheckComplete(isHuman) {
  if (!isHuman) {
    alert("Please correctly identify the cats.");
  }
  else {
    passThroughFormSubmit = true;
    formElt = document.getElementById("mainForm");
    formElt.submit();
  }
}
</script>

<h1>ASIRRA</h1>

<p>Today, the most common HIPs ask users to identify text that has been distorted or obscured, like the example seen to the right. Unfortunately, such challenges can be difficult and frustrating for people, yet are often easily solved by computers.</p>

<p>Asirra (Animal Species Image Recognition for Restricting Access) is a HIP that works by asking users to identify photographs of cats and dogs. This task is difficult for computers, but our user studies have shown that people can accomplish it quickly and accurately. Many even think it's fun!</p>

<form method="post" onsubmit="return MySubmitForm();" id="mainForm">
  <fieldset>
    <legend>Sample form</legend>

    <label>Test field:</label><input type="text" />

    <script
src="http://challenge.asirra.com/js/AsirraClientSide.js"></script>
    <script>
asirraState.SetCellsPerRow(4);
    </script>

    <input type="submit" value="Υποβολή" />

  </fieldset>
</form>

```

5.3.4.3 Επεξεργασία απάντησης

Σε αντίθεση με τα υπόλοιπα εργαλεία που χρησιμοποιήσαμε το Asirra δεν παρέχει έτοιμο κώδικα για τον έλεγχο της απάντησης. Παρόλα αυτά δίνεται ένα υπόδειγμα κώδικα το οποίο ελέγχει την απάντηση.

```
<?php

function ValidateAsirraChallenge()
{
    global $passed;

    $AsirraServiceUrl = "http://challenge.asirra.com/cgi/Asirra";
    $ticket = $_GET['Asirra_Ticket'];

    $url =
$AsirraServiceUrl."?action=ValidateTicket&ticket=".$ticket;

    $ch = curl_init();
    curl_setopt($ch, CURLOPT_URL, $url);
    curl_setopt($ch, CURLOPT_RETURNTRANSFER, 1);
    curl_setopt($ch, CURLOPT_CONNECTTIMEOUT, 5);
    $resultXml = curl_exec($ch);
    curl_close($ch);

    $xml_parser = xml_parser_create();
    xml_set_element_handler($xml_parser, "startElement",
"endElement");
    xml_set_character_data_handler($xml_parser, "characterData");
    xml_parse($xml_parser, $resultXml, 1);
    xml_parser_free($xml_parser);

    if (!$passed)
    {
        die("Asirra validation failed!");
    }
}

?>
```

5.3.4.4 Αξιολόγηση

Το Asirra έχει το ξεκάθαρο πλεονέκτημα ότι δεν παρουσιάζει στον χρήστη παραμορφωμένες εικόνες αλλά δείχνει κανονικές εικόνες ζώων. Έτσι η αναγνώρισή τους είναι αρκετά πιο εύκολη. Επιπλέον σε κάθε ένα από αυτά υπάρχει και ένας σύνδεσμος που παραπέμπει στην σελίδα της εταιρείας Petfinder, από όπου ο χρήστης μπορεί να υιοθετήσει το ζώο αυτό, παρέχοντας και μία κοινωνική υπηρεσία.

Ο χρήστης πρέπει να έχει μία μικρή γνώση αγγλικών ώστε να μπορεί να αναγνωρίσει την ερώτηση. Αυτό μπορεί να είναι πρόβλημα για όσους δεν γνωρίζουν καλά αγγλικά καθώς δεν παρέχεται κανενός είδους πολυγλωσσική υποστήριξη ώστε να αλλάξει το κείμενο.

Ακόμα δεν προσφέρονται καθόλου δυνατότητες παραμετροποίησης της εμφάνισης του τεστ.

Η εγκατάσταση του δεν είναι απλή, γιατί αν και δίνονται λεπτομερείς οδηγίες για την εισαγωγή της JavaScript στη φόρμα δεν δίνεται κάποια βιβλιοθήκη που μπορεί να χρησιμοποιηθεί άμεσα. Το παράδειγμα κώδικα που δίνεται είναι αρκετά πολύπλοκο και μπορεί να αποτελέσει πρόβλημα για κάποιους προγραμματιστές που δεν είναι αρκετά έμπειροι.

Τέλος το γεγονός ότι οι εικόνες δεν φορτώνονται από τον server αλλά από κάποιον τρίτο διακομιστή είναι θετικό γιατί μειώνει τον φόρτο εργασίας, όμως όπως και όλα τα άλλα captcha αναγνώρισης εικόνων έχει το μειονέκτημα πως ο χρήστης θα πρέπει εν τέλει να κατεβάσει τις εικόνες, καθυστερώντας τη φόρτωση της σελίδας.

5.3.5 Sweetcaptcha

5.3.5.1 Γενικά

Το Sweetcaptcha είναι ένα διαδραστικό τεστ captcha το οποίο ζητάει από τον χρήστη να σύρει κάποια εικόνα πάνω σε μία άλλη.



Εικόνα 5-17 Μία φόρμα που χρησιμοποιεί το Sweetcaptcha.

Παρέχεται παράδειγμα κώδικα PHP, ενώ είναι διαθέσιμο και σαν άρθρωμα για τα συστήματα διαχείρισης περιεχομένου Wordpress και Joomla.

5.3.5.2 Εγκατάσταση

Για την εγκατάσταση του χρειάζεται αρχικά να ανοιχθεί ένας λογαριασμός στην σελίδα και να δηλωθεί το site που θα το χρησιμοποιεί.

sweet CAPTCHA

Your SweetCaptcha is ready !

Please fill in your site details

website:

email:

re-type email:

language: English

category: Random

This will reflect on the image style. For example:
if you choose books as your category, than all images will be books related etc.

[CONTINUE](#)

[Tour](#) [Contact us](#) [Get it now !](#) [FAQ](#) [Designers](#)

WordPress Joomla! PHP Android Works with iPhone

©2010 all right reserved to sweetcaptcha.com ltd. [Pin it](#) [Like](#) 165

Εικόνα 5-18 Για την χρήση του Sweetcaptcha είναι απαραίτητη η δημιουργία λογαριασμού.

Για την χρήση του δεν παρέχεται κάποια βιβλιοθήκη αλλά παρέχεται κώδικας php, ο οποίος χρειάζεται να τροποποιηθεί και να γραφτεί στο site.

```
<?php
/*
 * Call your Sweetcaptcha object here.
 */

define('APP_ID', 11017);
define('SWEETCAPTCHA_KEY',
'9e19d537ae7fb954b3b7c7488153f04a419f6917') ;
define('SWEETCAPTCHA_SECRET',
'ccd9dde3c7fa6acb499d63f4fc1d4a836d953b73');
define('SWEETCAPTCHA_PUBLIC_URL', 'sweetcaptcha.php');

$sweetcaptcha = new Sweetcaptcha(
    APP_ID,
    SWEETCAPTCHA_KEY,
    SWEETCAPTCHA_SECRET,
```

```

SWEETCAPTCHA_PUBLIC_URL
);

/*
 * Do not change below here.
 */

/**
 * Handles remote negotiation with Sweetcaptcha.com.
 *
 * @version 1.0.8
 * @since December 14th, 2010
 */

if (isset($_POST['ajax']) and $method = $_POST['ajax']) {
    echo $sweetcaptcha->$method(isset($_POST['params']) ?
$_POST['params'] : array());
}

class Sweetcaptcha {

    private $appid;
    private $key;
    private $secret;
    private $path;

    const API_URL = 'www.sweetcaptcha.com';

    function __construct($appid, $key, $secret, $path) {
        $this->appid = $appid;
        $this->key = $key;
        $this->secret = $secret;
        $this->path = $path;
    }

    private function api($method, $params) {

        $basic = array(
            'method' => $method,
            'appid' => $this->appid,
            'key' => $this->key,
            'secret' => $this->secret,
            'path' => $this->path,
            'is_mobile' => preg_match('/mobile/i',
$_SERVER['HTTP_USER_AGENT']) ? 'true' : 'false',
            'user_ip' => $_SERVER['REMOTE_ADDR'],
        );

        return $this->call(array_merge(isset($params[0]) ? $params[0]
: $params, $basic));
    }

    private function call($params) {
        $param_data = "";
        foreach ($params as $param_name => $param_value) {
            $param_data .= urlencode($param_name) . '=' .
urlencode($param_value) . '&';
        }
    }
}

```



```
if ( !($fs = fsockopen(self::API_URL, 80, $errno, $errstr,
10) ) ) {
    die ("Couldn't connect to server");
}

$req = "POST /api.php HTTP/1.0\r\n";
$req .= "Host: www.sweetcaptcha.com\r\n";
$req .= "Content-Type: application/x-www-form-
urldata\r\n";
$req .= "Referer: " . $_SERVER['HTTP_HOST'] . "\r\n";
$req .= "Content-Length: " . strlen($param_data) .
"\r\n\r\n";
$req .= $param_data;

$response = '';
fwrite($fs, $req);

while ( !feof($fs) ) {
    $response .= fgets($fs, 1160);
}

fclose($fs);

$response = explode("\r\n\r\n", $response, 2);

return $response[1];
}

public function __call($method, $params) {
    return $this->api($method, $params);
}
?>
```

Αυτό το αρχείο χρειάζεται να συμπεριληφθεί στο αρχείο της σελίδας.

```
<?php require_once 'include/sweetcaptcha.php'; ?>
```

Επιπλέον η εφαρμογή κάνει χρήση της βιβλιοθήκης JavaScript JQuery η οποία πρέπει και αυτή να φορτωθεί.

```
<script type="text/javascript"
src="https://ajax.googleapis.com/ajax/libs/jquery/1.4.2/jquery.min.js"
"> </script>
```

Για να εισαχθεί αυτό το αρχείο στη σελίδα πρέπει να εισάγουμε τον παρακάτω κώδικα.

```
<?php echo $sweetcaptcha->get_html(); ?>
```

5.3.5.3 Επεξεργασία απάντησης

Για την επεξεργασία της απάντησης χρειάζεται να δημιουργηθεί ένας πίνακας με κάποιες από τις τιμές που έχει στείλει η φόρμα:

```
$scValues = array('sckey' => $_POST['sckey'], 'scvalue' =>
$_POST['scvalue'], 'scvalue2' => $_POST['scvalue2']);
```

Έπειτα χρειάζεται να καλεστεί η μέθοδος check του αντικειμένου sweetcaptcha.

```
$sweetcaptcha->check($scValues)
```

Ανάλογα με την τιμή επιστροφής της απάντησης προβάλλουμε και το αντίστοιχο μήνυμα. Ο πλήρης κώδικας της φόρμας είναι ο εξής:

```
<?php include("include/header.php"); ?>

<?php require_once 'include/sweetcaptcha.php'; ?>

<script type="text/javascript"
src="https://ajax.googleapis.com/ajax/libs/jquery/1.4.2/jquery.min.js
"></script>

<h1>Secureimage</h1>

<p>Το Sweetcaptcha είναι ένα διαδραστικό τεστ captcha το οποίο ζητάει
από τον χρήστη να σύρει κάποια εικόνα πάνω σε μία άλλη.</p>
<p>Παρέχεται σαν βιβλιοθήκη PHP, ενώ είναι διαθέσιμο και σαν άρθρωμα
για τα συστήματα διαχείρισης περιεχομένου Wordpress και Joomla.</p>
<?php if(empty($_POST)){ ?>
<form method="post">
  <fieldset>
    <legend>Sample form</legend>

    <label>Test field:</label><input type="text" />

    <?php echo $sweetcaptcha->get_html(); ?>

    <input type="submit" value="Υποβολή" />

  </fieldset>
</form>
<?php
}else{
  $scValues = array('sckey' => $_POST['sckey'], 'scvalue' =>
$_POST['scvalue'], 'scvalue2' => $_POST['scvalue2']);
  if ($sweetcaptcha->check($scValues) != "true") {
```

```
        echo "<br /><h3>The security code entered was incorrect." ;
        echo "Please go <a href='javascript:history.go(-1)''>back</a>
and try again.</h3>" ;
    }else{
        echo "<br /><h3>The security code entered was correct.</h3>" ;
    }
}
?>

<?php include("include/footer.php"); ?>
```

5.3.5.4 Αξιολόγηση

Η εφαρμογή Sweetcartcha διαφέρει πάρα πολύ σε σχέση με τις υπόλοιπες καθώς είναι αρκετά πρωτότυπη. Το γεγονός ότι οι χρήστες πιθανότατα δεν έχουν συναντήσει κάτι παρόμοιο μπορεί να το κάνει αρκετά ευχάριστο στην επίλυση. Επιπλέον το να σύρει ο χρήστης την εικόνα είναι πιο σύντομο από το να πληκτρολογήσει μία απάντηση.

Στον τομέα της χρήσης από άτομα με ειδικές ανάγκες υπάρχει και το επιπλέον πρόβλημα ότι, εκτός του ότι δεν μπορεί να χρησιμοποιηθεί από άτομα με προβλήματα όρασης, δεν μπορεί να χρησιμοποιηθεί και από άτομα με κινητικά προβλήματα, τα οποία μπορεί να βρουν την διαδικασία του να σύρουν κάτι αρκετά δύσκολο. Ακόμη δεν προσφέρεται κάποια εναλλακτική δοκιμασία ήχου.

Η ερώτηση που γίνεται είναι πάντα στα αγγλικά και είναι μέτριας δυσκολίας, οπότε μη αγγλόφωνες μπορεί να έχουν πρόβλημα. Δεν υπάρχει κάποια δυνατότητα επιλογής γλώσσας. Ακόμη δεν υπάρχει καμία πρόβλεψη για την παραμετροποίηση της εμφάνισης της εφαρμογής.

Τέλος καθώς χρειάζεται να είναι ενεργοποιημένη η γλώσσα JavaScript για να λειτουργήσει μπορεί να μην μπορεί να χρησιμοποιηθεί από χρήστες που την έχουν απενεργοποιημένη.

5.3.6 Civil Rights CAPTCHA

5.3.6.1 Γενικά

Το Civil Rights CAPTCHA είναι ένα νέο σχετικά τεστ, το οποίο συνδυάζει ερωτήσεις κειμένου με παραμόρφωση χαρακτήρων. Ο τρόπος λειτουργίας του είναι ότι παραθέτει μία είδηση που είναι σχετική με κάποια παραβίαση των ανθρωπίνων δικαιωμάτων με καθαρό κείμενο. Έπειτα έχει τρία διαφορετικά συναισθήματα, δύο θετικά και ένα αρνητικό με πορφή παραμορφωμένων λέξεων και ρωτάει τον χρήστη ποιο συναίσθημα του δημιουργεί η είδηση.



Εικόνα 5-19 Η φόρμα μας κάνοντας χρήση του Civil Rights CAPTCHA.

Πρόκειται για το πιο κοινωνικά ευαισθητοποιημένο captcha που συναντήσαμε, καθώς ο σκοπός του είναι να ευαισθητοποιήσει τους χρήστες για τις παραβιάσεις των ανθρωπίνων δικαιωμάτων.

Παρέχεται σαν βιβλιοθήκη για την PHP, ή σαν άρθρωμα για το σύστημα διαχείρισης περιεχομένου Drupal και σαν πρόσθετο για σελίδες που χρησιμοποιούν τις τεχνολογίες ColdFusion και Node.js.

5.3.6.2 Εγκατάσταση

Για την εγκατάστασή του χρειάζεται να κατεβάσουμε την βιβλιοθήκη PHP από την διεύθυνση <http://code.google.com/p/civil-rights-defenders-captcha/downloads/>. Περιλαμβάνονται δύο αρχεία, ίδια η βιβλιοθήκη και ένα υπόδειγμα κώδικα. Αντιγράψαμε την βιβλιοθήκη στον φάκελο includes.

Για την χρήση του χρειάζεται να συμπεριλάβουμε το αρχείο της βιβλιοθήκης στην φόρμα.

```
require_once("include/CRCphplib.php");
```

Έπειτα, στο σημείο που θέλουμε να εισαχθεί χρειάζεται να γράψουμε τον εξής κώδικα PHP.

```
<?php  
$crccaptcha = new civilrightscaptcha();  
$crccaptcha->show();  
?>
```

5.3.6.3 Επεξεργασία απάντησης

Τέλος για την επεξεργασία της απάντησης χρειάζεται να καλέσουμε την μέθοδο check του αντικειμένου civilrightscaptcha με παράμετρο δύο τιμές που παραδίδει η φόρμα.

```
$crccaptcha->check($_POST['crc_captcha'], $_POST['crc_sessid'])
```

Ανάλογα με την τιμή επιστροφής παρουσιάζουμε και το κατάλληλο μήνυμα.

Ο πλήρης κώδικας παρουσιάζεται παρακάτω:

```
<?php
    include("include/header.php");
    require_once("include/CRCphpplib.php");
?>

<h1>Civil Rights CAPTCHA</h1>

<p>With over 200 million CAPTCHAs being solved everyday, we hope that
by catching a tiny amount of those interactions we can help promote
and empower our partners - brave human rights defenders, who often
put themselves at great risk through their engagement for other
people's rights. The application has been developed by the
organization Civil Rights Defenders who's purpose is to work with
local organisations in longterm partnerships, striving to be present
in countries and regions where human rights are least respected.</p>
<?php if(!isset($_REQUEST['crc_captcha'])){ ?>
<form method="post">
    <fieldset>
        <legend>Sample form</legend>

        <label>Test field:</label><input type="text" />

        <?php
            $crccaptcha = new civilrightscaptcha();
            $crccaptcha->show();
        ?>

        <input type="submit" value="Υποβολή" />

    </fieldset>
</form>
<?php
}else{
    $crccaptcha = new civilrightscaptcha();
    if($crccaptcha->check($_POST['crc_captcha'],
$_POST['crc_sessid'])) {
        echo "<br /><h3>The security code entered was incorrect.";
        echo "Please go <a href='javascript:history.go(-1)''>back</a>
and try again.</h3>";
    }
    else {
        echo "<br /><h3>The security code entered was correct.</h3>";
    }
}
}
```

```
?>  
<?php include("include/footer.php"); ?>
```

5.3.6.4 Αξιολόγηση

Το μεγαλύτερο πρόβλημα του Civil Rights CAPTCHA είναι η επιλογή του περιεχομένου του. Αυτό συμβαίνει γιατί η εφαρμογή θέτει κάποια ηθικά ερωτήματα που μπορεί να μην έχουν την ίδια απάντηση για όλους τους χρήστες, όπως για παράδειγμα τα δικαιώματα των ομοφυλοφίλων, οι οποίοι μπορεί να βρουν το θέμα προσβλητικό. Επιπλέον κάποια άλλα θέματα μπορεί να χαρακτηριστούν ακατάλληλα για παιδιά.

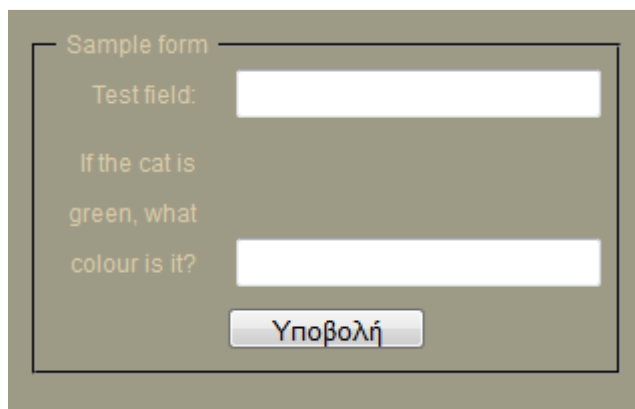
Επιπλέον οι δυνατές απαντήσεις που μπορεί να δώσει ο χρήστης, καθώς είναι κάποια συναισθήματα, είναι περιορισμένες, καθιστώντας εύκολη την δημιουργία μίας βιβλιοθήκης που θα τις περιέχει κατηγοριοποιημένες και κάνοντας πολύ πιο εύκολο για ένα πρόγραμμα να αναγνωρίσει τις λέξεις. Από εκεί και πέρα μπορεί να μην καταβληθεί καμία προσπάθεια για την κατανόηση της είδησης και απλά να χρησιμοποιηθεί η μοναδική λέξη με θετική ή αρνητική σημασία.

Πέρα από τα παραπάνω η εφαρμογή δεν έχει πολυγλωσσική υποστήριξη και καθώς οι ειδήσεις και οι απαντήσεις είναι στα αγγλικά οι χρήστες πρέπει να έχουν καλή γνώση της γλώσσας. Επιπλέον δεν δίνονται δυνατότητες διαμόρφωσης της εμφάνισής της.

5.3.7 Text CAPTCHA

5.3.7.1 Γενικά

Το Text CAPTCHA είναι μία δοκιμασία CAPTCHA που χρησιμοποιεί ερωτήσεις κειμένου για να ελέγξει αν ο χρήστης είναι άνθρωπος. Πρόκειται για μία από τις παλιότερες εφαρμογές captcha. Η σελίδα της εφαρμογής ισχυρίζεται ότι περιέχει πάνω από 180 εκατομμύρια ερωτήσεις.



Εικόνα 5-20 Μία φόρμα με μία ερώτηση του Text CAPTCHA.

5.3.7.2 Εγκατάσταση

Πριν τη χρήση της εφαρμογής χρειάζεται να γίνει αίτηση για την παροχή ενός δωρεάν κλειδιού, το οποίο είναι απαραίτητο.

Register for the textCAPTCHA service



If you want to experiment with the service to see if it will be suitable, no need to signup immediately. Using the API key **demo** will retrieve a limited subset of questions.

We require some personal information from you so we can contact you with service announcements. Rest assured that we value your privacy: we will only contact you when absolutely necessary, and will never share you details with any 3rd party.

Firstname

Surname

Email

Which of thirty three, 7, eighty six or 72 is the biggest?

Register for a new account

Forgotten/Lost your API Key?

Email

Remind me of my API key

textCAPTCHA.com

180,243,205 questions.
All easy.

[Really?](#)

[Why \(Not?\)](#)

[Demo](#)

[How it works](#)

[Sign Up](#) |

[API](#)

[Status](#)

Εικόνα 5-21 Η φόρμα εγγραφής στην υπηρεσία Text CAPTCHA.

Καθώς δεν παρέχεται κάποια βιβλιοθήκη πρέπει να εισαχθεί ολόκληρος ο κώδικας στη σελίδα. Αυτό γίνεται με τις παρακάτω εντολές:

```
$url = 'http://api.textcaptcha.com/demo';
try {
    $xml = @new SimpleXMLElement($url,null,true);
} catch (Exception $e) {
    // if there is a problem, use static fallback..
    $fallback = '<captcha>'.
        '<question>Is ice hot or cold?</question>'.
        '<answer>'.md5('cold').'</answer></captcha>';
    $xml = new SimpleXMLElement($fallback);
}
// display question as part of form
$question = (string) $xml->question;
// store answers in session
$ans = array();
foreach ($xml->answer as $hash)
    $ans[] = (string) $hash;
$_SESSION['captcha'] = $ans;
```

Η ερώτηση αποθηκεύεται στην μεταβλητή question η οποία χρησιμοποιείται για την προβολή της.

```
<label><?php echo $question; ?></label><input type="text"
name="captcha_code"/>
```

5.3.7.3 Επεξεργασία απάντησης

Η απάντηση παίρνεται από τα στοιχεία που αποστέλλει η φόρμα με κάποια επεξεργασία.

```
$ans = $_POST['captcha_code'];
$ans = strtolower(trim($ans));
$ans = md5($ans);
```

Μετά ελέγχεται αν αυτή έχει ήδη αποθηκευτεί στο session κατά την δημιουργία της φόρμας.

```
in_array($ans, $_SESSION['captcha'])
```

Ανάλογα με το αν περιέχεται ή όχι παράγεται το κατάλληλο μήνυμα. Ο πλήρης κώδικας της φόρμας είναι ο εξής:

```
<?php include("include/header.php"); ?>

<h1>Text CAPTCHA</h1>

<p>Το Text CAPTCHA είναι μία δοκιμασία CAPTCHA που χρησιμοποιεί
ερωτήσεις κειμένου για να ελέγξει αν ο χρήστης είναι άνθρωπος.
Πρόκειται για μία από τις παλιότερες εφαρμογές captcha.</p>
<?php if(!isset($_REQUEST['captcha_code'])) {

    // load captcha using web service
    $url = 'http://api.textcaptcha.com/demo';
    try {
        $xml = @new SimpleXMLElement($url, null, true);
    } catch (Exception $e) {
        // if there is a problem, use static fallback..
        $fallback = '<captcha>'.
            '<question>Is ice hot or cold?</question>'.
            '<answer>'.md5('cold').'</answer></captcha>';
        $xml = new SimpleXMLElement($fallback);
    }
    // display question as part of form
    $question = (string) $xml->question;
    // store answers in session
    $ans = array();
    foreach ($xml->answer as $hash)
        $ans[] = (string) $hash;
    $_SESSION['captcha'] = $ans;
```



```
?>
<form method="post">
  <fieldset>
    <legend>Sample form</legend>

    <label>Test field:</label><input type="text" />

    <label><?php echo $question; ?></label><input type="text"
name="captcha_code"/>

    <input type="submit" value="Υποβολή" />

  </fieldset>
</form>
<?php
}else{
  $ans = $_POST['captcha_code'];
  $ans = strtolower(trim($ans));
  $ans = md5($ans);

  if (!in_array($ans,$_SESSION['captcha'])) {
    echo "<br /><h3>The security code entered was incorrect.";
    echo "Please go <a href='javascript:history.go(-1)''>back</a>
and try again.</h3>";
  }else{
    echo "<br /><h3>The security code entered was correct.</h3>";
  }

  unset($_SESSION['captcha']);
}
?>

<?php include("include/footer.php"); ?>
```

6 Συμπεράσματα

Η ανάλυση των τεχνικών CAPTCHA μας άφησε με αρκετά αρνητικά συναισθήματα, τα οποία θα παρουσιάσουμε παρακάτω.

Αρχικά, πέρα από τις ιδιαιτερότητες κάθε τεχνικής υπάρχουν κάποια θέματα που επηρεάζουν γενικά τις σελίδες που χρησιμοποιούν τέτοιου είδους τεστ. Το κυριότερο είναι πως οι χρήστες που θα κληθούν να συμπληρώσουν μία φόρμα μπορεί να αποφύγουν να το κάνουν αν δουν ένα τεστ τέτοιου είδους ή ακόμη περισσότερο αν την συμπληρώσουν και ανακαλύψουν πως έκαναν λάθος στο CAPTCHA. Αυτό σημαίνει χαμένη κυκλοφορία για τη σελίδα, το οποίο στις περισσότερες περιπτώσεις συνεπάγεται οικονομικές απώλειες.

Πέρα από το παραπάνω γενικό συμπέρασμα τα τεστ CAPTCHA που υπάρχουν σήμερα έχουν αρκετά μειονεκτήματα. Αν και ο αρχικός σκοπός μας ήταν να κατατάξουμε τις διαφορετικές τεχνικές, με την ανάλυσή τους έγινε φανερό πως όλες μοιράζονται αρκετά προβλήματα.

Το μεγαλύτερο πρόβλημα είναι πως ο σκοπός οποιουδήποτε τεστ είναι εξ' ορισμού να θέσει ένα εμπόδιο στη συμπλήρωση μιας φόρμας, μειώνοντας της ευχρηστία της σελίδας και κατά συνέπεια το conversion rate, δηλαδή το ποσοστό των χρηστών που συμπληρώνουν τη φόρμα. Από τη στιγμή που ο χρήστης πρέπει να καταβάλει επιπλέον προσπάθεια για να απαντήσει στην ερώτηση αυτόματα θα προσπαθεί να το αποφύγει.

Συγκεκριμένα είδη, κυρίως τα cartcha παραμόρφωσης γραμμάτων και τα ακουστικά cartcha είναι πιο εύκολο να σπάσουν καθώς οι αντίστοιχες τεχνολογίες αυτόματης επίλυσής τους έχουν αναπτυχθεί αρκετά. Σαν αντίμετρο η δυσκολία των ερωτήσεων αυξάνεται συνεχώς, κάνοντας τους χρήστες να τα συνδυάζουν με μία δυσάρεστη εμπειρία που θέλουν να αποφύγουν. Οι σελίδες θα μπορούσαν να παρουσιάζουν διάφορα εναλλακτικά τεστ και ο χρήστης να συμπληρώνει μόνο ένα από αυτά, όμως αυτό σημαίνει επιπλέον κόστος κατά την ανάπτυξη της σελίδας γιατί θα χρειαστεί να υλοποιηθούν παραπάνω από μία δοκιμασίες.

Επιπλέον όλα τα είδη δεν μπορούν να χρησιμοποιηθούν από συγκεκριμένες κατηγορίες ατόμων με ειδικές ανάγκες, αν και αυτό το πρόβλημα μπορεί να λυθεί δίνοντας στον χρήστη την δυνατότητα να επιλέξει ένα τεστ διαφορετικού είδους. Για παράδειγμα είναι σχετικά διαδεδομένη πρακτική στα cartcha παραμόρφωσης γραμμάτων να συνδυάζονται και με ένα εναλλακτικό ακουστικό cartcha.

Το μοναδικό είδος τεστ που φαίνεται να ξεχωρίζει κατά την εξέτασή μας είναι τα cartcha αναγνώρισης εικόνας, γιατί είναι αρκετά εύκολα στην επίλυση και πάρα πολύ δύσκολο να σπάσουν. Όμως η υλοποίησή τους είναι επίσης δύσκολη και δεν είναι επιλύσιμα από χρήστες με προβλήματα όρασης.

Ειδικά για τις υπηρεσίες που χρησιμοποιήσαμε στην εφαρμογή που υλοποιήθηκε, κάθε μία είχε μία διαφορετική δυσκολία υλοποίησης και διαφορετική ποιότητα στο αποτέλεσμα. Δεν ξεχωρίζει κάποια από αυτές καθώς όλες έχουν διαφορετικά πλεονεκτήματα και μειονεκτήματα.

Κάποια κοινά μειονεκτήματα που μοιράζονται οι περισσότερες είναι η απουσία δυνατότητας παραμετροποίησης της εμφάνισής τους και η απουσία πολυγλωσσικής υποστήριξης, καθώς οι περισσότερες απευθύνονται σε αγγλόφωνες.

Βέβαια αξίζει να σημειώσουμε ότι τα cartcha χρησιμοποιούνται για να επιλύσουν ένα πολύ έντονο πρόβλημα, το οποίο αν μία σελίδα δεν προσπαθήσει να το ελέγξει μπορεί να την βγάλει εκτός λειτουργίας ή να μειώσει την ποιότητα του περιεχομένου της, οπότε δεν μπορούμε να αποκηρύξουμε την χρήση τους.

Με βάση τα παραπάνω και την κατάσταση της τεχνολογίας τη στιγμή που γράφεται αυτή η εργασία θα προβούμε σε δύο προτάσεις. Η αναγνώριση εικόνων ξεχωρίζει σαν η πιο αποτελεσματική τεχνική η οποία είναι και η λιγότερο δυσάρεστη για τον χρήστη. Η υλοποίησή της είναι πιο δύσκολη από τα υπόλοιπα αλλά τα πλεονεκτήματα που προσφέρει δικαιολογούν το επιπλέον κόστος. Δυστυχώς αυτό το είδος cartcha δεν μπορεί να χρησιμοποιηθεί από άτομα με προβλήματα όρασης.

Για αυτή την ομάδα των χρηστών θα πρέπει να παρέχεται και μία εναλλακτική δοκιμασία. Το είδος της δεν είναι τόσο εύκολο να το προτείνουμε, καθώς οι δοκιμασίες που είναι προσβάσιμες σε άτομα με προβλήματα όρασης είναι οι ακουστικές και οι δοκιμασίες κειμένου, και με τις δύο κατηγορίες να έχουν σαφή και έντονα μειονεκτήματα. Δεν μπορούμε να κάνουμε μία σαφή πρόταση για την υλοποίηση κάποιας από αυτές, καθώς θα πρέπει κατά περίπτωση να επιλέγεται η καλύτερη εναλλακτική με τέτοιο τρόπο ώστε να καλύπτει καλύτερα τις ανάγκες κάθε εφαρμογής.

7 Βιβλιογραφία

- (n.d.). Ανάκτηση από Cryptographp : <http://www.captcha.fr/>
- Ahn, L. v., Blum, M., Hopper, N. J., & Langford, J. (n.d.). *CAPTCHA: Using Hard AI Problems for Security*.
- Are You Human? CAPTCHA's Many Ways of Asking the Same Question* . (n.d.). Ανάκτηση από techi.com: <http://www.tech.com/2010/05/are-you-human-captchas-many-ways-of-asking-the-same-question/>
- ASIRRA. (n.d.). Ανάκτηση από Microsoft Research: <http://research.microsoft.com/en-us/um/redmond/projects/asirra/>
- Bursztein, E., Bethard, S., Fabry, C., Mitchell, J. C., & Jurafsky, D. (n.d.). *How Good are Humans at Solving CAPTCHAs? A Large Scale Evaluation*.
- CIVIL RIGHTS CAPTCHA . (n.d.). Ανάκτηση από civilrightsdefenders.org: <http://captcha.civilrightsdefenders.org/>
- Egele, M., Bilge, L., Kirda, E., & Kruegel, C. (2010). *CAPTCHA Smuggling: Hijacking Web Browsing Sessions to Create CAPTCHA Farms*.
- esti, D. L. (2005). *Leveraging the CAPTCHA Problem*.
- Holman, J., Lazar, J., Heidi, J., & D'Arcy, J. (2007). *Developing Usable CAPTCHAs for Blind Users*.
- JavaScript. (n.d.). Ανάκτηση από Wikipedia: <http://en.wikipedia.org/wiki/JavaScript>
- Jeremy Elson, J. R., & Saul, J. (2007). *Asirra: A CAPTCHA that Exploits Interest-Aligned Manual Image Categorization*.
- Mori, G., & Malik, J. (n.d.). *Recognizing Objects in Adversarial Clutter: Breaking a Visual CAPTCHA*.
- PHP. (n.d.). Ανάκτηση από Wikipedia: <http://en.wikipedia.org/wiki/Php>
- PHP: Hypertext Preprocessor. (n.d.). Ανάκτηση από PHP.net: <http://php.net/>
- reCAPTCHA. (n.d.). Ανάκτηση από wikipedia.org: <http://en.wikipedia.org/wiki/ReCAPTCHA>
- reCAPTCHA: Stop Spam, Read Books. (n.d.). Ανάκτηση από google.com: <http://www.google.com/recaptcha>
- Schlaikjer, A. (2007). *A Dual-Use Speech CAPTCHA: Aiding Visually Impaired Web Users*.
- Securimage. (n.d.). Ανάκτηση από phpcaptcha.org: <http://www.phpcaptcha.org/>
- Shirali-Shahreza, M. H., & Shirali-Shahreza, M. (2007). *Multilingual CAPTCHA* .

Text CAPTCHA Logic Questions. (n.d.). Ανάκτηση από textcaptcha.com:
<http://textcaptcha.com/>

Ximenes, P., Santos, A. d., Fernandez, M., & Jr., J. C. (2006). *A CAPTCHA in the Text Domain.*