KV Web Engineering
SS 2010

durchgeführt am
Institut für Anwendungsorientierte Wissensverarbeitung
der Johannes Kepler Universität Linz

Leitung: a.Univ.-Prof. Dr. Birgit Pröll

Aufgabe 2/a

Gruppe F

26.05. 2010

The Inaccessibility of CAPTCHA - Alternatives to Visual Turing Tests on the Web & Breaking a Visual CAPTCHA

Kurze Einführung:

Die wissenschaftlichen Publikationen, die wir ausgewählt haben, gehören zum Thema Accessibility. (In der Aufarbeitung wurde häufig die englischen Ausdrücken benutzt, hoffentlich ist es nicht unklar). In diesem Thema Accessibility haben wir uns die CAPTCHA gewählt und die zwei Publikation stammen aus der W3C WebSeite. Die erste Publikation ist gerade erreichbar von der W3C Seite und die Zweite befindet sich in der Referenzliste der ersten Publikation.
Wir haben den Titel der Publikationen behalten, aber unsere Ausarbeitung konzentriert mehr auf „Accessiblity of Captcha" und andere Lösungen mit Vorteilen und Nachteilen anstatt des Captcha. (Die verwendeten Dokumente wurden erstmals von englisch auf ungarisch danach auf deutsch übersetzt).

About Captcha & The Problem:

Heuzutage existiert eine gemeinsame Methode zur Beschränkung des Zugangs zu Dienstleistungen über das Internet überall. Das heisst visuelle Überprüfung von Bitmap-Bildern (Visual Verification of a Bitmapped Image). Das grösste Problem gibt es für Nutzer, die blind oder eine Sehbehinderung haben oder irgendeine andere Art von Behinderung wie Legasthenie. Dieser Artikel probiert einige Lösungen, welches ein System enthält, welches die Zugänglichkeit für behinderte Menschen ermöglicht. Dieses, der am weitesten verbreitete Lösung ist heuzutage die grafische Darstellung des Textes, die oft bei Registrierung oder Eintragung benutzt wird. Diese Seiten versuchen herauszufinden, ob die Ressourcen (oder die Dienstleistungen) ein richtiger Mann oder eine Web-Roboter nutzen möchte. Der Benutzer muss einen verzerrten Text in einem Bitmap-Bild lesen, danach in eine Form geben. Dieses Verfahren der Pioniere der Carnegie Mellon Universität, die es Captcha (Completely Automated Public Turing Test Computers and Humans Apart Tell) nannten. Viele Gruppen arbeiten an diesem oder ähnlichen Verfahren, die zusammen auch Captcha genannt werden. Zu diesem gehört ein Test, den der Turing-Test heisst. Der Name wurde nach einem berühmten Mathematiker Alan Turing benannt. Das Ziel ist mit diesem Test, den Mann von einem Computer zu unterscheiden.

Für diese Art der visuellen und textlichen Prüfung zahlen die Nutzern ein hohen Preis, die blind oder sehbehindert sind. In vielen Fällen machen die Seiten es unmöglich für Menschen mit Behinderungen einen „account" herzustellen, zu kommentieren, oder nur etwas über das Internet zu kaufen. Zusammenfassend, die Captchas sind für Menschen mit Behinderungen **durchgefallen** zu benutzen.

Zunächst wurden die Captchas bei größeren Seiten verwendet, da man ihre Diensleistungen leicht missbrauchen könnte („spams" Zielen, illegale Aktivitäten). In späteren Zeiten wurde es wie ein

inaccessible-by-design-technology unter den kleinen Seiten verbreitete und die Technologie ist wirklich vollständig gemischt worden. Zum Beispiel: Bank ING „PIN Guard" ist eine visuelle Tastatur, die die Buchstaben mit dem Benutzer des „Passcode" vergleicht.
Die Captchas wird oft für persönliches Blogs verwendet. Viele Blogger benutzen erfolgreich um die Kommentar-Spam zu verhindern, aber diese Captchas sind oft einfach und können sehr einfach umgangen werden. Es gibt viele andere Methoden, die besser anwendbar wären, und damit wäre die Seite auch erreichbar für Menschen mit Behinderungen. (Am Ende dieses Artikel werden wir auch mehr Typen von Captchas und Ergebnisse von „cracked Captchas" vorstellen.)

The Possible Solutions:

Viele Technologien sind für den Benutzer zur Verfügung, das sie verhindern, einen Account zu machen oder zu benutzen. Viele von ihnen sind mindestens genauso wirksam wie die visuelle Prüfung, während sie mehr verfügbar für Menschen mit Behinderungen sind. Unten gibt's viele verwendbare Alternative zusammen mit Vorteilen und Nachteilen (pros and cons):

Logic Puzzles

Der Zweck der Überprüfung der Visualisierung ist die Menschen von den Maschinen zu unterscheiden. Die Idee ist die menschliche Logik zu benutzen. Einfache Mathematik, Word-Puzzles, usw.
Pros:

- einfach zu implementieren
- menschliche Stärke
- Schwäche von Computern

Cons:

- Die Menschen mit kognitiven Behinderungen haben noch immer Schwierigkeiten
- Leuten, die die Fremdsprache nicht sprechen
- Die Antworten sollten flexibel gehandhabt werden und das System sollte ein grosse Datenbank mit Fragen und Antworten gehalten werden. Die Antworten sollten immer weiterverschoben werden, so dass die Spiders nicht alles (Antworten und Fragen) „besorgen" können.

Sounds Output

Anstatt der visuellen Überprüfung oder neben der visuellen Überprüfung
Pros:

- Es ist eine Alternative

Cons:

- Benutzer sollten es mehrmals hören
- Nicht so einfach es zu verstehen (es ist oft mit anderen Lauten gemischt)
- keine Lösung für Taube
- Es gibt sehr viele Spracherkennungprogramme, die effizienter sind als das Imaging-Programm

Credit-Card Validation

Erfassung und Prüfung der Kreditkartennummer.
Pros:

- Die Karte repräsentiert die Identität
- Es ist leich erhältlich

Cons:

- Security issues
- Perception

Limited-use accounts

Es gibt Benutzern, die die Resourcen nur selten und nicht gleich erreichen sollen. Zum Beispiel: Die Benutzern, die Konzert-Tickets suchen, befristet auf drei, suchen einen Tag. Viele Seiten finden etwas heraus, dass die Häufigkeit der Interaktion limitiert. (Explizit – ein Konto deaktivieren für den Rest des Tages; Implizit - in der dramatischen Zunahme der Reaktionszeit).
Pros:
Dazu braucht es kein „user-checking"
Cons:
Es schliesst auch die legitimen Benutzer aus

Single sign-on

Ein System wird verwendet wie ein Passport to, bietet bereits etablierten Identitäten Zugang.
Pros:

- Es ist nicht mehr wichtig der Benutzer-Test

Cons:

- Passport kann auch Captchas benutzen.
- Privacy issues

Public-key infrastructure solutions

In diesem Fall kann eine Seite „certificate" benutzen und so kann sich der Benutzer selbst identifizieren.
Pros:
Der Benutzer-Test ist nicht wichtig in der Zukunft
Cons:
Es existiert noch nicht
Die Menschen mit Behinderungen

Biometrics

Es ist noch für Zukunft: Retina-untersuchung, fingerprints, Gesichgeometrie,

Pros:

- Abhängig von körperlichen Merkmalen

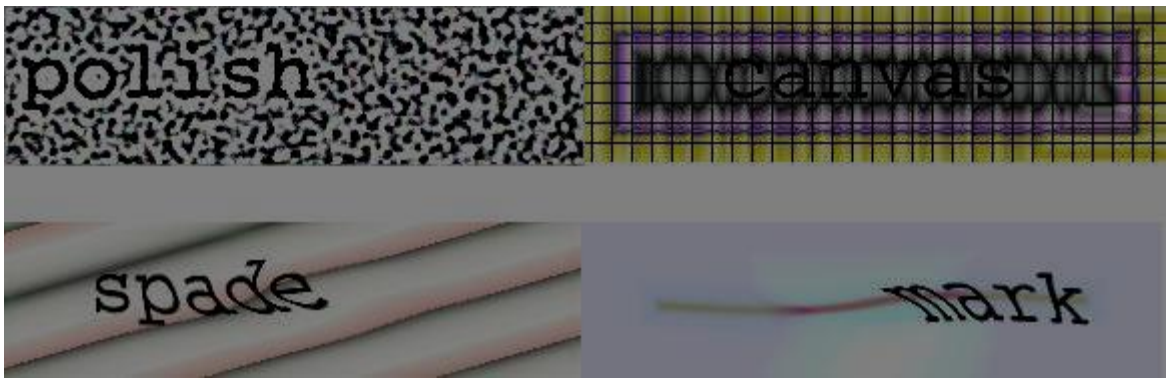
Cons:

- Es braucht noch dafür Zeit, dass diese Method verbreiten sollte.
- Die Menschen, die ohne Augen geboren ist oder Verletzung bei Fingern können diese Method nicht benutzen

Results about breaking captchas:

Hiermit möchten wir uns einigen Ergebnisse kurz vorstellen, einigen Beispiel von Captchas aufführen und zeigen, dass die Captchas mit einer entsprechenden (Bilderkennenung)Alghoritmen in einem guten Prozent ziemlich einfach aufbrechbar sind. Der Name des Program,das wurde benutzt, ist Gimpy. Das Programm sollte nur eingen Sekunden das Bild aufarbeiten und den (Klart)Text herausgeben. In diesem Experiment wurde 191 Bilden benutzt, davon der richtiger Text – in 92% Prozent – festgestellt geworden ist.

Einigen Beispielen, die (oben) erfolgreich und (unten) nicht erfolgreich aufgebrochen wurde:



Conlusion:

Die Web-Seiten, den millionen Benutzer haben, brauchen immer ein „System-Kontroll" ,dass die Missbrauchen minimalzieren können. In diesem Niveau sollte empfehlenswert paralelle Methoden verwenden. (Zum Beispiel: Audio und visuelle Captcha zusammen) . Wir können es auch markieren, dass menschlichen Benutzern immer durch diese Systeme (auf)brechen zu werden, deswegen brauchen die Seiten immer etwas Bestätigung von Menschen. Das bedeutet ,dass die Menschen mit Behinderungen immer mit der Maschinen (durch Interaktion) kommunizieren sollen, damit ihrer menschlichen Existenz zu bestätigen.

Auf der anderen Seite, die „low-resources" Seiten können die Captchas manchmal überflüssig verwenden. Sogar dann, wenn die Qualität des Captchas wirklich schlecht ist (Früher könnten wir es feststellen, dass man viele Arten von Captchas manchmal ohne Probleme aufbrechen kann.)
Für diese Seiten wäre empfehlenwert anstatt der Captchas anderen Methoden zu benutzen.

Quellen:

- http://www.w3.org/TR/turingtest/
- http://www.cs.sfu.ca/~mori/research/gimpy/
- http://www.w3.org/2004/Talks/0319-csun-m3m/slide23-0.html