Implementering och utvärdering av en bildbaserad autentiseringsmetod

MAGNUS BERGMAN
och VIKTOR GUMMESSON

KTH Datavetenskap
och kommunikation

Examensarbete
Stockholm, Sverige 2011

Implementering och utvärdering av en bildbaserad autentiseringsmetod

M A G N U S B E R G M A N
o c h V I K T O R G U M M E S S O N

Examensarbete i datalogi om 15 högskolepoäng
vid Programmet för datateknik
Kungliga Tekniska Högskolan år 2011
Handledare på CSC var Mikael Goldmann
Examinator var Mads Dam

URL: www.csc.kth.se/utbildning/kandidatexjobb/datateknik/2011/bergman_magnus_OCH_gummesson_viktor_K11076.pdf

Kungliga tekniska högskolan
Skolan för datavetenskap och kommunikation

KTH CSC
100 44 Stockholm

URL: www.kth.se/csc

Förord

Implementering och utvärdering av en bildbaserad autentiseringsmetod

För att autentisera sig på Internet idag finns det en de facto standard som tillämpas, vilket är alfanumeriska lösenord. Denna metod är enkel att använda både ur ett användar- och ett utvecklarperspektiv, men det finns en växande oro över säkerheten.

Det forskars om en del alternativa metoder för autentisering och några av dem fokuseras på användarens förmåga att minnas och identifiera bilder. I rapporten kommer en bildbaserad autentiseringsmetod beskrivas och implementeras, en mindre studie på denna implementation kommer även att redovisas. Resultaten var intressanta men vi fick inte in tillräckligt med data för att dra någon statistiskt säkerhetsställd slutsats.

Abstract

Implement and evaluate a cognitive authentication method

When it comes to authenticating a user on the Internet today, the de facto standard is to do so with an alphanumeric password. This method is easy to use both from a user- and from a developers prespective. However there are growing concerns about the security of this scheme.

There exists researches about alternatives to the alphanumeric scheme, some of these focus on the user's ability to remember and identifying pictures. Such a scheme are reserched and implemented in this thesis and a small study about the useability of such a scheme is presented. The results, though intresting, were to slim to do any statistically secure conclusion.

Arbetsfördelning

Detta arbete har bestått utav två delar, en implementation och rapport. På implementationen har Viktor haft ansvar och gjort framsidan, inloggningen och registreringen där Magnus har hjälpt till när det behövts. Magnus gjorde undersökningen och den sista biten vilket var att lägga upp sidan och den gjorda databasen på en server och se till att publicera hela sidan.

Rapporten är skriven av oss båda men eftersom Viktor hade större ansvar av implementationen har Viktor haft ansvar för kapitel två och fyra och Magnus har haft ansvar för förorden, kapitel ett, tre och fem.

Innehållsförteckning

1 Introduktion

Ämnet för denna rapport är kognitiv autentiseringsmetod. Vilket är ett sätt att autentisera en användare, vilket skiljer sig från de klassiska alfanumeriska lösenorden. I dessa metoder använder man bilder som lösenord istället. Fördelen med detta är säkerheten då det antas vara mycket svårare att avlyssna användares lösenord.

1.1 Syfte

Behovet av att kunna identifiera personer på ett säkert sätt har funnits väldigt länge. Ute i värden har vi baserat våra identitetskontroller på ett personligt möte, där en jämförelse av signatur och bild ur en identitetshandling. Men när vi skall identifiera oss på Internet kan inte denna metod användas. Den de facto standard som används istället är då en kombination av ett användarnamn och ett hemligt alfanumeriskt lösenord.

På senare tid har en oro över säkerheten av denna inloggningsmetod ökat. Detta på grund utav studier som gjorts, vilka visar att väldigt många användare har samma lösenord och att de allra flesta är osäkra. Dessa lösenord är också väldigt utsatta och svaga för avlyssningsattacker som vi diskuterar mer i 1.1.2.

I följande avhandling ämnar vi att undersöka om det går att ersätta de klassiska alfanumeriska lösenorden med kognitiv autentiseringsmetod ur ett användarvänlighetsperspektiv. För att kunna avgöra detta kommer en implementation av ett inloggningssystem som använder just bilder som lösenord göras. Därefter skall en liten studie på hur väl användare kunde använda systemet när dem precis fått sitt lösenord genomföras.

1.2 Bakgrund

1.2.1 Autentisering på Internet idag

På Internet idag så används alfanumeriska lösenord för identifiering. När man registrerar sig på en hemsida får man lämna uppgifter om

sig själv och sätta ett lösenord. Därefter när man vill identifiera sig kommer hemsidan att jämföra de nu lämnade uppgifterna mot det som är lagrat från tillfället man registrerade sig. Huruvida dessa stämmer överens avgör om man blir identifierad eller ej.

1.2.2 Oro för säkerheten

Det finns ett stort problem med de alfanumeriska lösenorden som används på Internet idag. Det är inte direkt i metoden själva problemet ligger i, utan det är i hur den används utav oss människor. Oftast får man som användare själv välja ett lösenord, och som användare vill man ha ett lösenord som är lätt att komma ihåg, vilket resulterar i ganska enkla lösenord för de flesta användarna.

Man kan läsa om detta i [IMP2010] som är en studie på de klassiska alfanumeriska lösenorden. Man har här kollat på 32 miljoner lösenord som olika användare har. Resultaten från undersökningen är signifikanta, de visar bland annat att 20% av alla användarna delar på cirka 5 000 lösenord och att 60% av alla lösenords längd är åtta eller färre tecken långa.

Dessa lösenord är också svaga på det sätt att det är lätt för en attakerare att avlyssna dessa. Det är båda den elektroniska avlyssningen och problemet med att folk kan tjuvkika på en användare som skriver in sitt lösenord. Eftersom dessa klassiska lösenord är statiska och inmatningen sker på samma sätt varje gång räcker det att se den exakta inmatningen en gång för att få tag på ett lösenord. De klassiska elektroniska attackerna på lösenord är bland annat en så kallade keylogger som installeras på offrets dator vilket sparar ner alla tangenttryckningar, men även paketavlyssnare som genskjuter offrets nätverkstrafik och försöker hitta lösenordet i paketen.

1.3 Kognitiv autentiseringsmetod

Ett alternativ till dessa klassiska lösenord är något som på engelska kallas kognitiv autentiseringsmetod. Dessa metoder lägger tyngden på människans förmåga att känna igen och memorera bilder. Denna typ av metod föreslogs först i [BLO1995] där användaren skulle bli identifierad genom att välja ut en delmängd punkter från en bild i en korrekt ordning.

Men idag finns det väldigt många olika metoder som bygger på samma princip. Ett exempel är [Passfaces] som går ut på att identifiera mänskliga ansikten. Ett annat exempel är att man skall hitta tre av sina lösenordsbilder i en stor mängd bilder och klicka inom den triangel dessa tre bilder bildar. Ett sista exempel är [PassPoint] som är likt det första förslaget, det gäller då att klicka på ett antal punkter i korrekt ordning, nackdelen med denna metod är att lösenordet blir lika statiskt som ett klassiskt alfanumeriskt lösenord.

Det har gjorts en del forskningar på ämnet tidigare, en av de är [CBO2007] som fått fram att deras metod var bra om man kollar på antal lyckade inloggningar. Därefter kom de fram till att valet av bilder påverkar hur lätt det är att logga in, detta motsäger tidigare forskning på ämnet. Men de visade också att hålla reda på flera grafiskt lösenord samtidigt kan bli svårt. Sen har vi [GW2007] som gör en kryptoanalys på dylika system och kommer fram till att inte alla bildbaserade system är säkra.

2 Implementationen

2.1 Beskrivning av implementationen

Vår implementation fungerar på följande sätt. En användare blir vid registreringen tilldelad ett lösenord som består utav fem bilder. Sedan autentiserar man sig genom att kunna känna igen och identifiera bilder från sitt lösenord ur en större mängd bilder.

Vid tillfället då man vill autentisera sig får användaren se en större mängd bilder i form utav en fem gånger fem matris, där en bild ur användarens lösenord kommer att finnas med. Det man sedan skall göra är att välja ut den rad i matrisen där bilden finns med. För att bli autentiserad måste man göra detta fem gånger och lyckas välja rätt rad alla fem gånger.

2.2 Implementation

Vi valde att implementera detta projekt som en hemsida. Eftersom det skulle utföras en mindre användarvänlighetsstudie skulle det bli lättare att nå ut till personer, och även lättare för dem att testa och utvärdera vår implementation. All information om användare, lösenord och loggar lagras i en databas. De 124 bilder som vi använder till matrisen och till lösenord lagras även i samma databas som binärdata. Utvecklingen har skett i ASP.net.

Vid förstasidan får man se en välkomsttext som förklarar innehållet och syftet med hemsidan. Man har tre valmöjligheter, logga in som existerande användare, göra en ny användare eller att utföra vår undersökning som är till för vår studie. Detta visas i bild 2.1.

Bild 2.1 Förstasidan

2.2.1 Registrering

Vid registreringen blir man tilldelad ett slumpmässigt genererat lösenord som man måste memorera. Man får samtidigt skriva in sitt namn, namnen är unika och det får inte finnas några dubbletter. Som visas i bild 2.2.

Bild 2.2 Registrering

2.2.2 Autentisering

När man har påbörjat sin inloggning möts man av en text som beskriver hur man skall gå tillväga och en fem gånger fem matris av unika bilder. I en av de fem rader finns den första bilden i användarens lösenord. Användaren blir ombedd att välja ut den rad där bilden finns. Detta upprepas fem gånger (alltså en gång för varje bild i lösenordet). Nedan visas ett exempel av en användare som håller på att logga in i bild 2.3.

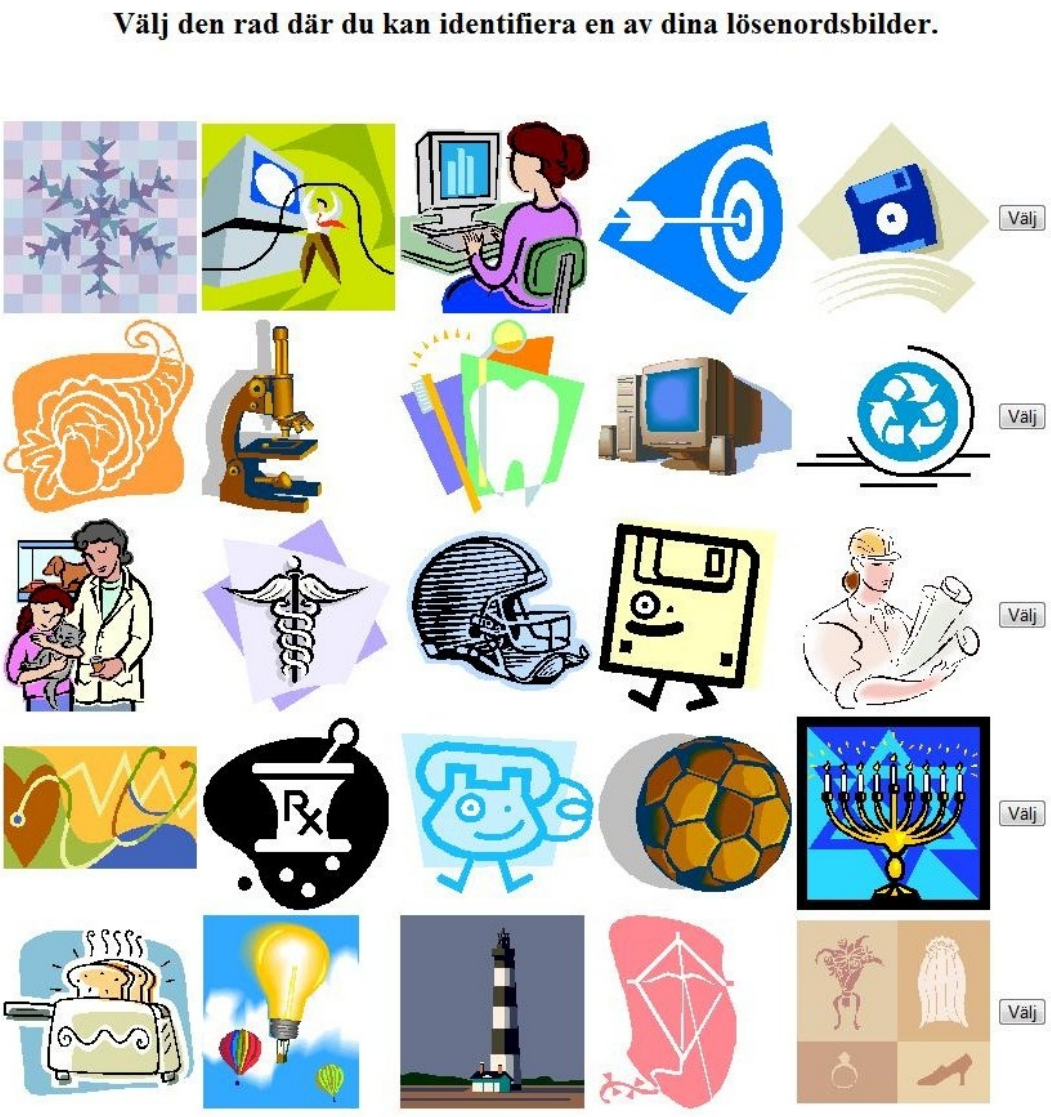

Bild 2.3 Autentisering

Sedan blir användaren vidareskickad till en sida som talar om huruvida användaren lyckades logga in eller ej. Sedan finns en länk som tar användaren tillbaka till förstasidan där den kan välja att försöka logga in igen, ta del av undersökningen, skapa ny användare eller välja att lämna hemsidan.

2.2.3 Undersökning

Undersökningen består utav fyra frågor, en med flersvarsmöjlighet och sedan tre där endast ett svar krävs. Vi ber användaren att fylla i frågorna för att få svar på vilka problem som användaren hade med metoden, om den skulle kunna tänka sig använda den här metoden, hur pass bra användaren skulle kunna bli på att fylla i rätt lösenord och svara på hur användarvänligt de upplevde att systemet var.

2.3 Begränsningar som vi satt

Vi har begränsat oss till att låta 124 bilder ligga i databasen, som vi använder till lösenord och att generera de fem gånger fem matriserna som används vid autentisering.

Vid registrering har användaren ej möjlighet att helt fritt väja innehåll eller storlek på sitt lösenord. Vi har valt att inte låta användaren använda fria bilder, alltså bilder som inte finns i vår databas. Användaren får i stället ett fem bilder långt lösenord med slumpmässigt valda bilder ur vår databas.

3 Användarvänlighetsstudien

En kort studie på användarvänligheten av vårt system gjordes. Fokuset låg på hur väl användarna lyckas logga in och komma ihåg sitt lösenord precis efter registreringen.

3.1 Metod

Vår studie sköttes över Internet på vår hemsida och vi träffade aldrig försökspersonerna. Vi lade upp en hemsida som beskrevs i kapitel två. På hemsidan fanns det instruktioner om hur vårt system fungerade och en länk till vår undersökning som också beskrivs i appendix A. Andra delen av studien var att kolla hur det gick med alla inloggningar, därför fördes en logg på alla inloggningsförsök och i den sparades om användaren lyckades logga in eller ej.

3.2 Resultat

Nedan visas resultaten om hur inloggningarna gick för användarna. Det var totalt nio personer som registrerade sig och det vara åtta personer som svarade på undersökningen. Varje gång någon av användarna försökte logga in i vårt system sparades resultatet av inloggningen., resultatet av detta visas i tabel 3.1.

Antal lyckade / antal försök	Procent
18/21	86%

Tabel 3.1 Resultat från inloggnings loggningen.

Vår undersökning bestod utav fyra stycken frågor om olika aspekter av vår inloggningsmetod som användarna fick testa. Frågorna och resultatet ser ni här nedan i tabel 3.2 och formuläret för undersökningen återfinns i appendix A.

Fråga 1(Flervalsfråga) - **Vilka av de följande meningarna beskriver de problem du hade med inloggningsmetoden?**

Det tog lång tid för bilderna att laddas.	0%
Det var svårt att memorera mitt lösenord.	75%
Det var svårt att lokalisera mina bilder.	38%

Fråga 2 - **Skulle du kunna tänka dig att använda denna metod för autentisera dig på Internet istället för traditionella lösenord?**

Ja. Det är en bättre lösning på alla sätt.	0%
Ja. Men bara på sidor med känslig information.	0%
Ja. Men bara på sidor som inte har känslig information.	88%
Nej.	12%

Fråga 3 - **Hur pass sannolikt är det att du skulle kunna träna in ditt lösenord så att du skulle kunna klara av att logga in 9 av 10 gånger?**

Inte sannolikt alls.	0%
2	0%
3	0%
4	50%
Högst sannolikt.	50%

Fråga 4 - **På en skala 1 till 5 hur pass användarvänligt tycker du att detta system var?**

Mycket dåligt.	0%
2	12%
3	50%
4	25%
Mycket bra.	13%

Tabel 3.2 Resultatet av undersökningen.

4 En kort säkerhetsstudie

Denna gjordes även fast den är väldigt liten är just för att dessa system börjar komma fram som alternativ eftersom man vill bli mer resistent mot just övervakningsattacker av olika slag, så det kändes relevant att göra ett mindre test även fast fokuset ligger på användarvänligheten.

4.1 Metod

Vi har utfört följande test där två personer har deltagit. Den ena har suttit och loggat in medan en andra suttit bredvid och försökt att lista ut den första personens lösenord. Efter varje inloggning fick tjuvkikaren testa att logga in och fick sedan reda på hur många rätt den hade. Detta upprepades tio gånger.

4.2 Resultat

Det visade sig vara väldigt svårt att få reda på ett lösenord genom att tjuvkika. Det närmaste en person var att få fram ett lösenord var två stycken bilder utav fem. Näst bästa var en bild utav fem.

5 Diskussion

Tyvärr måste vi medge att det var svårt att hitta villiga testare och fick därför för lite med data för att kunna dra några statistiskt säkerhetsställda slutsatser. Att vi inte fick ihop tillräckligt med testare beror först och främst på den mycket begränsade budgeten som vi hade till detta projektet, både i tid och pengar.

Vi valde också på grund utav tidsbrist att göra vissa begränsningar på vår implementation. Det vi valde att göra var att sätta alla lösenord till fem bilder och vi valde att göra matrisen fem gånger fem bilder stor. Detta på grund utav att vi redan innan visste att det skulle vara svårt att få in tillräckligt med testdata och inte skulle ha data nog för att kunna se någon skillnad mellan olika långa lösenord med mera.

Man kan i alla fall från vår studie se att de som testade vårt system hade det ganska lätt att lyckas logga in, hela 86% av fallen lyckades de med inloggningen. Detta kan bero på många saker, en av dem är att vi valde att ha ett relativt kort lösenord. Men vår studie var också bara precis efter registrering och vi kan inte dra några slutsatser om hur bra man kommer ihåg sitt lösenord efter en längre tid. Men direkt efter registrering verkar det vara lätt att logga in, som man även ser i undersökningen tror de flesta att med lite träning skulle de kunna lära sig att lyckas logga in nio gånger utav tio.

Det var lite spridda tankar från våra testare om användarvänligheten. I vår undersökning kan man då se att hälften ger vårt system en trea på en skala från ett till fem i användarvänlighet och resterande är jämnt fördelade runt det. Detta tyder på att de traditionella lösenorden är lite lättare att använda än vårt system.

Man kan också från undersökningen se att de som testade detta system är oroliga över säkerheten. Vi anser att den oron är berättigad, i alla fall när man har de begränsningar vi satt på systemet. Risken för att en som bara testar ett lösenord helt slumpmässigt och lyckas logga in är 1/(5^5) = 0.00032(fem valmöjligheter och fem rundor) vilket är en liten för hög siffra i en inloggningsmetod. Men en fördel med vårt system är att när man testat en kombination av rader kan man inte förkasta den i nästa försök eftersom vårt system bygger på ett slumpmässigt element som gör att rätt svar ändras hela tiden. Det som måste göras för att få ner numret är att öka lösenordslängden eller att öka antalet rader i matrisen där man skall identifiera sin lösenordsbild.

Vi testade även systemet mot en tjuvkikare som kollar över axeln på den som skriver in sitt lösenord. Det visade sig vara väldigt svårt, om man testade att replikera lösenordet bara efter man hade kollat på en inloggning lyckades ingen få ett enda rätt av dem fem lösenordsbilderna. Trots att vi fick se en person logga in tio gånger på en kort tid var det bästa resultatet att lyckas identifiera två utav fem bilder. Det som är svårt är att man bara får några få sekunder på sig på varje matris med tjugofem bilder och ännu mindre tid att se vilken rad som personen sedan väljer och försöka memorera fem bilder, och lika snabbt börjar det om och du måste göra det hela fem gånger under tio till tjugo sekunder. Vi kom fram till att detta system är säkert mot tjuvkikare, man behöver inte oroa sig över sin omgivning när man ska logga in med denna metod.

Ur ett användarvänlighetsperspektiv ser det bra ut, men vi kan själva inte dra några slutsatser om det definitivt är lika bra som våra siffror visar men det tyder helt klart på att mer forskning inom området bör utföras för att kunna avgöra om denna metod har en framtid eller inte. Det finns många områden man måste utföra mer forskning på.

5.1 Framtidsstudier

För framtida projekt är det viktigt att se till att göra en större studie för att få statistiskt säkerhetsställd data för att kunna dra ordentliga slutsatser. Man borde även lägga ner mer tid på att göra en ordentlig säkerhetsstudie. Man skulle också göra en kombinerad studie för att se ungefär vart man skall lägga sig för att få bra användarvänlighet och bra säkerhet.

Andra punkter som är viktiga att fokusera på är hur bra man kommer ihåg sitt lösenord en längre tid, säg en vecka eller en månad. Även hur bildvalen påverkar en användares förmåga att komma ihåg sitt lösenord och även om det är en bättre ide att låta användaren välja bilder till sitt eget lösenord.

Referenser

[IMP2010] Imperva
Consumer Password Worst Practices, 2010
Senast besökt: 2011-03-27
http://www.imperva.com/docs/WP_Consumer_Password_Worst_Practices.pdf

[BLO1995] G. E. Blonder
Graphical Password
US Patent 5559961, filed 1995-08-30
Senast besökt: 2011-04-05
http://www.freepatentsonline.com/5559961.html

[Passfaces] Passfaces
Senast besökt: 2011-04-05
http://www.realuser.com

[PassPoint] Sonia Chiasson, Alain Forget, Elizabeth Stobert, P. C. van Oorschot and Robert Biddle.
Multiple password interference in text passwords and click-based graphical passwords. In CCS '09: Proceedings of the 16th ACM conference on Computer and communications security, pages 500–511, New York, NY, USA, 2009. ACM.

[CBO2007] S. Chiasson, R. Biddle, P.C. van Oorschot
A Second Look at the Usability of Click-Based Graphical Passwords
Carleton University, 2007
Senast besökt: 2011-04-07
http://hotsoft.carleton.ca/~sonia/content/Chiasson_SOUPS2007_Click_based_GP.pdf

[GW2007] P. Golle, D. Wagner
Cryptanalysis of a Cognitive Authentication Scheme

IEEE Symposium on Security and Privacy 2007: 66-70
Senast besökt: 2011-04-06
http://crypto.stanford.edu/~pgolle/papers/sat.pdf

Appendix A

Fråga 1(Flervalsfråga) - **Vilka av de följande meningarna beskriver de problem du hade med inloggningsmetoden?**

- Det tog lång tid för bilderna att laddas.
- Det var svårt att memorera mitt lösenord.
- Det var svårt att lokalisera mina bilder.

Fråga 2 - **Skulle du kunna tänka dig att använda denna metod för autentisera dig på Internet istället för traditionella lösenord?**

- Ja. Det är en bättre lösning på alla sätt.
- Ja. Men bara på sidor med känslig information.
- Ja. Men bara på sidor som inte har känslig information.
- Nej.

Fråga 3 - **Hur pass sannolikt är det att du skulle kunna träna in ditt lösenord så att du skulle kunna klara av att logga in 9 av 10 gånger?**

- Inte sannolikt alls.
- 2
- 3
- 4
- Högst sannolikt.

Fråga 4 - **På en skala 1 till 5 hur pass användarvänligt tycker du att detta system var?**

- Mycket dåligt.
- 2
- 3
- 4
- Mycket bra.