Hjemmeeksamen i IN2120 H2019

T63: PET (Privacy Enhancing Technology) teknikker og trender

Studentnr.: 614610

Innledning

Det kommer stadige nye teknologier som tas i bruk, og spesielt de siste årene har man sett hvordan ny teknologi også kan føre med seg nye risikoer for brudd på personvernet og datasikkerheten. Man ser oftere overskrifter om sikkerhetsbrudd hos ulike virksomheter i nyhetene, og mange har blitt mer bevisst på overvåkning av nettaktivitet og kommunikasjonen man foretar seg over nettet, som følge av at stadig mer personlig informasjon ligger tilgjengelig på nett. Men betyr det at utviklingen av ny teknologi kun vil føre med seg en økt svekkelse av personvernet?

Det å sikre personvernet i teknologien har vært et tema innenfor informasjonssikkerhet i lang tid. Begreper som dataminimering, med formål om å begrense mengden personopplysninger som benyttes, anonymisering og pseudonymisering for å begrense identitetsmarkører, har blant annet vært viktige verktøy for å styrke personvernet. Med utgangspunkt i disse teknologiene oppstod etter hvert betegnelsen PET (Privacy Enhancing Technologies/Personvernøkende teknologi), som skal favne hele spekteret av teknologier og organisatoriske tiltak som bidrar til å øke personvernet og datasikkerheten. [1]

Jeg vil i oppgaven gå nærmere inn på hva PETs er, gi noen eksempler på aktører innen PET og eksisterende teknologi, og se på utfordringer man står ovenfor når det gjelder å beskytte personvernet. I tillegg vil jeg stille spørsmål ved om vi kan stole på at PET-teknologiene i tilstrekkelig grad vil kunne beskytte personopplysningene våre i fremtiden.

Hoveddel

Privacy Enhancing Technologies (PETs), direkte oversatt til «personvernsøkende teknologier» er en betegnelse på teknologier som har som formål å minske risikoen for brudd på datasikkerheten og personvernet. PETs forsøker å gjøre dette gjennom å eliminere eller redusere mengden personopplysninger som tas i bruk, og forhindre at unødvendig og uønsket behandling av personopplysninger finner sted, samtidig som man unngår å miste funksjonaliteten til informasjonssystemet.

Personvernsøkende teknologier skiller seg fra andre områder innen informasjonssikkerhet på den måten at PETs først og fremst har som formål å beskytte personopplysninger og identitet. PET skiller seg også fra tekniske tiltak som brannmurer, antivirus, spamfilter og annen programvare, ved at de omfatter tekniske *organisatoriske* tiltak for å beskytte personvernet. Målene med PET-teknologien er å kunne:

- Beskytte personopplysninger og forsikre de som bruker teknologien om at opplysningene deres blir behandlet konfidensielt.
- Gi brukeren av teknologien økt kontroll over personlig data som blir brukt av nettjenester.
- Minimere personlig data som blir samlet inn og brukt.
- Tilby anonymitet til brukeren gjennom pseudonymisering og anonymisering av data.
- Jobbe mot at brukeren skal kunne gi et informert samtykke til nettsidetilbyderen ved bruk av personopplysninger.
- Gjøre det vanskeligere å bryte regler for datasikkerhet og bidra til å oppdage brudd.

- Hjelpe organisasjoner med å sikre beskyttelse av data og å overholde den generelle databeskyttelsesforordningen (GDPR). [2] [1]

Bakgrunn

Bakgrunnen for at begrepet «Privay enhancing technology» oppstod, var at det ble brukt i forbindelse med rapporten; «Privacy-enhancing technologies: the path to anonymity» som ble lagt frem av datatilsynsmyndighetene i Ontario, Canada og Nederland i 1995. Rapporten ble senere presentert under den 17. internasjonale datatilsynsmyndighetskonferansen i København samme år.
I rapporten viste man hvordan de fleste informasjonssystemer kan designes på en måte som gjør at brukerens identitet kan holdes delvis eller helt skjult. For å kunne gjøre dette trengte man en koblingssentral hvor personidentifiserende identifikatorer, for eksempel navn eller fødselsnummer, blir gjort om til pseudonymer, slik at de ikke direkte kan knyttes til en identifiserbar person.
Selve ideen om anonymitet kan imidlertid sies å ha oppstått flere år tidligere, i 1981, da David Chaum publiserte en artikkel om hvordan man kunne oppnå anonymitet for avsender og mottaker av elektronisk kommunikasjon. Før dette hadde man i kryptografien kun fokusert på hvordan man kunne skjule innholdet i data fra uautorisert innsyn, slik at kun de med tilgang definert av en tilgangspolicy skulle ha innsyn i innholdet. Det at man kunne skjule hvem som kommuniserte med hvem, og ikke bare hva som ble kommunisert, var derfor et stort gjennombrudd i forhold til å kunne beskytte personvernet uten å gå på bekostning av funksjonalitet. [3]

PETs i Europa

Det meste av utviklingen av Personvernsøkende teknologi i dag foregår i Europa, og en viktig aktør innen PET er The European Union Agency for Cybersecurity (ENISA). ENISA holder til i Athen og Heraklion i Hellas, og har jobbet med å bedre cybersikkerheten i Europa siden 2004. ENISA samarbeid tett med eksperter på personvern, og finansierer og følger med på forskning og utvikling av PETs, blant annet utført av IPEN (Internet Privacy Engineering Network). De holder også en årlig konferanse kalt Annual Privacy Forum, som er et viktig samlingssted for industri, forskning og beslutningstakere til å diskutere personvern. [4]

PETs i GDPR

PETs er blant annet nevnt i GDPR, General Data Protection Regulation, eller Personvernsforordningen, i Artikkel 25: *Innebygd personvern og personvern som standardinnstilling*.
I artikkelen kreves det at den *behandlingsansvarlige*, som er ansvarlige for håndtering av personopplysninger, skal gjennomføre egnede tekniske og organisatoriske tiltak både i forbindelse med behandling, og under behandling av personopplysninger. Denne formen for å tenke personvern i alle ledd kalles *innebygd personvern*.
Eksempler på personvernstiltak som nevnes i GDPR er *pseudonymisering*, hvor personopplysninger blir behandler på en måte som sørger for at de ikke kan knyttes til en

identifiserbar person, og *dataminimering*, som innebærer å begrense mengden personopplysninger som benyttes til det mest nødvendige.

GDPR sier at disse tiltakene er ment å sikre at «personopplysninger som standard ikke gjøres tilgjengelige for et ubegrenset antall fysiske personer uten den berørte personens medvirkning.» [5]
For å sjekke at tiltakene blir fulgt, er det vanlig for virksomheter å gjennomføre egne internkontroller.

Hva er en personopplysning?

Hovedformålet med PETs å begrense mengden personopplysninger som benyttes. En personopplysning blir av datatilsynet definert som: «Alle opplysninger og vurderinger som kan knyttes til deg som enkeltperson». Eksempler på personopplysninger er: navn, adresse, telefonnummer, e-post og fødselsnummer. I tillegg er bilder av gjenkjennelige personer, dynamiske IP-adresser, lydopptak og biometri, som fingeravtrykk, irismønster og hodeform, og adferdsmønster, både på nett og i det virkelige liv, alle regnet som personopplysninger.

De 12 kategoriene av personopplysninger som regnes som spesielt *sensitive* er:

Opplysninger om rasemessig eller etnisk opprinnelse, opplysninger om politisk oppfatning, opplysninger om religion, opplysninger om filosofisk overbevisning, opplysninger om fagforeningsmedlemskap, genetiske opplysninger, biometriske opplysninger, helseopplysninger, opplysninger om seksuelle forhold, opplysninger om seksuell legning, opplysninger om straffedommer og opplysninger om lovovertredelser. [6]

Eksempler på personvernsøkende teknologi:

Så hvordan kan PET bidra til å beskytte disse personopplysningene? Det finnes mange forskjellige Personvernsøkende teknologier tilgjengelig for internettbrukere i dag, med ulike funksjoner. Noen tillater brukeren å kontrollere informasjonskapsler som blir lagret på harddisken, og andre tillater å surfe anonymt på nettet. Personvernsøkende teknologier sørger generelt for at brukeren kan være på nett samtidig som han/hun får vernet om sine personlige opplysninger. Eksempler på noen PET-teknologier som finnes og brukes i utstrakt grad i dag er:

Pseudonymisering:

Pseudonymisering innebærer at data blir gjort uidentifiserbare, gjennom at informasjon som kan gjøre det mulig å identifisere et individ, kalt identifikatorer, blir byttet ut med en eller flere kunstige identitetsmarkører, eller *pseudonymer*. Et pseudonym som erstatter for data, gjør den opprinnelige dataen mindre identifiserbar, samtidig som dataen fortsatt kan brukes til å for eksempel gjennomføre analyser. [7]

Obfuskering:

En måte å maskere/skjule data for uautoriserte, ved å legge til forstyrrende og misvisende elementer, som for eksempel tegn eller annen data. Obfuskeringen kan gjørs manuelt eller ved

hjelp av et program, og ulike teknikker innebærer blant annet å navngi variablene med forvirrende navn, eller å få koden til å se annerledes ut gjennom å endre syntaks og form. Et eksempel hentet fra en internasjonal konkurranse i obfuskering, International Obfuscated C Code Contest, viser en kode omformet til en labyrint gjennom å legge til mellomrom [11]:

```
char*M,A,Z,E=40,J[40],T[40];main(C){for(*J=A=scanf(M="%d",&C);
--              E;              J[              E]              =T
[E   ]=  E)   printf("._");  for(;(A-=Z=!Z)  ||  (printf("\n|"
)    ,   A    =              39              ,C              --
)    ;   Z    ||    printf   (M   ))M[Z]=Z[A-(E   =A[J-Z])&&!C
&    A   ==              T[                                  A]
|6<<27<rand()||!C&!Z?J[T[E]=T[A]]=E,J[T[A]=A-Z]=A,"_.":" |"];}
```

Fordeler med obfuskering er at den gjør det både vanskelig og tidkrevende å hente ut dataene som er obfuskert, den sikrer dataene fra uautorisert tilgang, og gjør det også mulig å forminske størrelsen på kildekoden. Ulemper med obfuskering er at den ikke er helt umulig å reversere, i tillegg til at enkelte antivirusprogram vil si ifra dersom man havner på en nettside som er obfuskert, siden obfuskering også blir brukt til å gjemme skadelig kode. [8]

Differensielt personvern:

Differensielt personvern er personvernsforbedrende teknikk som brukes av blant annet Google og Apple, der brukerens personlige opplysninger blir tillagt tilfeldig informasjon, såkalt «statistisk støy», før de analyseres. På denne måten kan ikke dataene direkte kobles til en bestemt person. Kun når opplysningene settes sammen med opplysninger fra mange andre brukere, kan man se mønstre og gjennomsnitt av dataene. På denne måten kan Apple for eksempel se hva mange av sine brukere søker på, eller hvilke emojier som blir brukt oftest, uten at de har tilgang til hva hver enkelt bruker skriver på tastaturet sitt.
Hos Apple foregår differensielt personvern som første steg i et dataanalyse-system, hvor personvernet blir ivaretatt i hvert steg:

1. I det første steget blir informasjonen gjort privat gjennom lokal differensielt personvern på brukeren sin enhet. Her blir dataene tilført «støy», slik at Apple sine servere ikke mottar rene data fra brukeren.

2. Videre blir identifikatorer knyttet til enheten fjernet fra den innsamlede dataen, og dataene blir overført til Apple via en kryptert kanal.

3. Apples analysesystem bearbeider de private dataene, og fjerner IP-adresser og annen metadata.

4. Dataene blir i siste steg slått sammen og delt med de relevante Apple-teamene.

Hele analyseprosessen foregår i tillegg i et miljø med begrenset tilgang, slik at de anonymiserte dataene aldri er bredt tilgjengelig for de ansatte i Apple. [9]

Anonyme proxy-servere:

En «Anonymiserer», eller en anonym proxy-server, skjuler identiteten og aktiviteten på nett, for eksempel e-postadressen eller IP-adressen, ved å erstatte den med en usporbar identitet, som en engangs-mail, en tilfeldig IP-adresse, eller et pseudonym. Serveren fungerer som et skjold mellom klient-pcen og internett, som kobler seg på nettet på vegne av brukeren og på den måten skjuler brukerens personlige informasjon.
En anonym server kan for eksempel brukes til å omgå sensur av nettinnhold, som finnes enkelte land, eller hvis man vil unngå målrettet reklame på nettet. [10]
En av de best kjente «anonymisererne» er Tor, «The Onion Router». Tor er et gratis verdensomspennende nettverk av rutere som enkeltindivider eller grupper kan bruke til å rute datatrafikken gjennom. Dataene sendes via tilfeldige rutere og på den måten forhindres nettsider fra å spore den opprinnelige avsenderen.

Tilgang til personlige data:

Retten til tilgang til personlige data gir brukeren retten til å be tjenestetilbyderen om å få en kopi av sine personlige data som er lagret om dem, og også velge om man ønsker å eller slette dataen. På den måten kan brukeren få innsikt i hvorfor og hvordan de innsamlet innsamlede dataene blir brukt, og selv kontrollere at det foregår på en lovlig måte.

Enhanced privacy ID (EPID):

Enhanced privacy ID er en forbedring av DAA-algoritmen som finnes i mange av Intel sine prosessorer. I motsetning til tradisjonelle algoritmer for digital signatur (f.eks. PKI), hvor hver enhet har en unik offentlig verifiseringsnøkkel og en unik privat signaturnøkkel, har DAA-algoritmen en felles offentlig verifiseringsnøkkel til mange unike private signaturnøkler. DAA støtter anonymitet ved at en enhet kan bevise for en ekstern part at den er et autentisk medlem av en gruppe, uten å måtte oppgi identitet. EPID tilbyr i tillegg funksjonalitet som gjør det mulig å tilbakekalle nøkler som er kompromittert eller stjålet og utestenge disse fra gruppen, kun basert på en signatur opprettet av nøkkelen, uten å noen gang få kjennskap til nøkkelens identitet.
Eksempler på bruk av EPID er for å bevise at en enhet er ekte ved oppdatering av programvare, eller ved forespørsel om tilgang til et system.

Utfordringer

En utfordring med å legge til rette for at tekniske løsninger skal kunne ta hensyn til personvern, er at mange daglig veksler mellom ulike «identiteter». En identitet kan for eksempel enten være en rolle, en brukerkonto eller et brukernavn. Siden utviklingen i samfunnet går mot en økning i elektronisk samhandling, blir det også stadig mer utfordrende å skulle tilpasse identifisering og autentisering til hver enkelt rolle og relasjon mellom disse.

En annen faktor som skaper utfordringer for personvernet, er at vi legger igjen stadig flere digitale spor. Siden 90-tallet har denne økningen vært enorm, og mye skyldes et økende krav om å identifisere seg i ulike sammenhenger på nett. Elektroniske spor er noe av det som

har preget personverndebattene mest de siste årene. En risiko med elektroniske spor er at utviklingen av moderne analyseverktøy har gjort det lettere å analysere data, slik at det også er mulig for uvedkommende å finne frem til enkeltpersoners bevegelsesmønstre, interesser og annen personidentifiserende informasjon basert på disse sporene.

En annen utfordring er at det ikke finnes så mange klare retningslinjer i dag for organisasjoner i bruk av PET-løsninger, slik at de av organisasjonene som ikke bruker løsningene tar en større personvernsrisiko, som det igjen er vanskelig for brukeren å vite om og beskytte seg mot.
I tillegg er det for eksempel blitt vanligere å dele informasjon om helseopplysninger med apper, som ikke nødvendigvis har det samme personvernet som helsetjenesten.

Disse utfordringene fører med seg spørsmål som om det vil bli mulig å ha kontroll på egne personopplysninger, eller om vi må innfinne oss med at det alltid vil være vanskelig å sikre seg helt.

PETs i fremtiden

Eksperter på personvern er enige om at anonymisering av data vil spille en stadig viktigere rolle i fremtiden. For at PET-teknologier skal vokse og tas i bruk i enda større grad fremover, er det nødvendig med økt bevissthet om at de finnes, sørge for en bedre brukervennlighet, og å gi bedre insentiver til organisasjoner og bedrifter for å tilby eller implementere PET-løsninger.

Det drives i dag en del forskning og utvikling, med bidrag fra blant annet ENISA, på nye PETs-teknologier, modenheten for PETs til å tas i bruk, og hvordan teknologiene kan implementeres. Det finnes også prosjekter som GDA Score Project, som sammenligner ulike anonymiseringsteknikker, slik at det blant annet skal bli enklere for brukerne å velge den teknologien som passer best for deres formål. [12]

Avslutning

PETs har vist seg som nyttige verktøy med flere bruksområder for ulike aktører, og de kan åpne opp for mange nye muligheter, blant annet å kunne analysere sensitive data uten å gå på kompromiss med personvernet. Fortsatt gjenstår det mye forskning og utvikling for å fullt ut realisere potensialet, og det gjenstår en del før PETs er distribuert til markedet på en måte som gjør at flere selskaper benytter seg av dem. For at dette skal skje må det skapes en større bevissthet rundt denne teknologien, både for enkeltpersoner og selskaper. Her spiller myndigheter og offentlige organer en viktig rolle både i å øke bevisstheten, men også å bidra med retningslinjer som kan føre til at disse løsningene implementeres i ny teknologi som skapes, og med finansiering av forsking og utvikling av ny personvernsøkende teknologi.

PETs virker som et lovende bidrag til at vi i fremtiden kanskje kan ha enda større tillit til at dataene våre behandles helt anonymt. Dersom utviklingen går i riktig retning, med økt bruk og et økt fokus på personvern i bedrifter, er det ikke umulig at de personvernøkende teknologiene skal klare å møte utfordringene som ny teknologi fører med seg.

Referanser

1. “Summary of privacy enhancing technologies – a survey of tools and techniques”. *Cranium*. Besøkt: 25. Oktober 2019.

2. «Privacy enhancing technologies». *Wikipedia*. Sist endret: 6. Oktober 2019. Besøkt: 25. Oktober 2019.

3. «Personvernøkende teknologi og identitetsforvaltning». Thomas Olsen. Besøkt: 1. November 2019.

4. «Privacy Enhancing Technologies» . *ENISA (European Union Agency for Cybersecurity)*. Besøkt: 25. Oktober 2019.

5. «Artikkel 25: Innebygd personvern og personvern som standardinnstilling». Jan Sandtrø, teknologiadvokat. Besøkt: 25. Oktober 2019.

6. «Hva er en personopplysning». *Datatilsynet*. Sist endret: 17.07.2019. Besøkt: 25. Oktober 2019.

7.”What are Privacy-Enhancing Technologies (PETs)?”. *Golden data (blogg)*. Besøkt: 25. Oktober 2019.

8. «Obfuscation (software)». *Wikipedia*. Sist oppdatert: 23. September 2019. Besøkt: 25. Oktober 2019.

9. «Differential Privacy Overview». *Apple*. Besøkt: 01. November 2019.

10. «Anonymizer». *Wikipedia.* Sist oppdatert: 18. September 2019. Besøkt: 2. November 2019.

11. Don Libes, *Obfuscated C and Other Mysteries*, John Wiley & Sons, 1993, pp 425. ISBN 0-471-57805-3

12. “GDA Score – About”. *GDA-Score*. Besøkt: 03. November 2019.