NTNU
Det skapende universitet

Innføring av EUs datalagringsdirektiv

Forslag til teknisk og administrativ løsning for innsamling og bruk av elektroniske kommunikasjonsdata

Marianne Hove

Master i kommunikasjonsteknologi
Oppgaven levert: Juni 2010
Hovedveileder: Svein Johan Knapskog, ITEM

Norges teknisk-naturvitenskapelige universitet
Institutt for telematikk

Oppgavetekst

Datalagringsdirektivet ble vedtatt av EU i 2006 og skal etter planen implementeres av alle medlemsland og EØS-land. Formålet med direktivet er å gi myndighetene bedre muligheter til å bekjempe alvorlig kriminalitet ved at opplysninger om tele- og nettrafikk lagres. Data som skal lagres er slike som er nødvendige for å spore og identifisere alle parter som kommuniserer via tekst, bilder eller lyd. Data lagres på formen dato, klokkeslett og varighet for telefonsamtaler, tidspunkt og varighet av nettoppkobling, IP-adresser og bruker-ID for internettadgang. Dataene skal lagres og oppbevares i minst seks måneder og maksimalt to år. Data som avdekker innholdet i kommunikasjonen kan ikke lagres med hjemmel i direktivet.

Denne hovedoppgaven vil med bakgrunn i direktivet vurdere mulige tekniske og administrative løsninger for innsamling og bruk av elektroniske kommunikasjonsdata, og vurdere mulig innvirkning på personvernet.

Oppgaven gitt: 15. januar 2010
Hovedveileder: Svein Johan Knapskog, ITEM

Forord

Utførelsen av dette arbeidet ble utført fra 15.januar til 1.juni 2010, ved Institutt for Telematikk; Fakultet for informasjonsteknologi, matematikk og elektronikk (IME); Norges Teknisk- Naturvitenskaplige Universitet (NTNU). Målet med dette arbeidet var å oppnå kunnskap rundt et gitt problem innen informasjonssikkerhet gjennom systematisk og metodisk arbeid. Med utgangspunkt i EUs Datalagringsdirektiv var målsetningen å foreslå en implementeringsløsning med fokus på personvern og informasjonssikkerhet dersom en slik lagringspraksis skulle bli aktuell også i Norge. Fokuset lå på aksesskontroll gjennom autentisering og autorisasjon, kryptering av og søk i (kryptert) lagret data samt revisjonslogger var hovedfokus.

Jeg vil takke min professor og veileder her på NTNU, Svein Johan Knapskog, for hans oppfølging og grundige tilbakemeldinger gjennom arbeidsprosessen. Jeg har satt stor pris på tilbakemeldinger både av teknisk og administrativ karakter gjennom våre ukentlige møter. Takk for at du delte kvalifiserte meninger og evalueringer vedrørende den adresserte problemstillingen, og også for forslag til forbedringer underveis i arbeidsprosessen.

Jeg vil også takke min bi-veileder, Svein Willassen. Han har delt sin kunnskap og erfaring vedrørende norsk lovverk og digital etterforskning med meg, og jeg føler mulighetene ved en god teknisk lagringsløsning har blitt fremhevet, selv om spørsmålet rundt en mulig innføring av direktivet og lovpålagt lagring har vært debattert heftig i media mens utførelsen av dette arbeidet har pågått. Selv om vår kommunikasjon i hovedsak har vært via e-post, har

jeg ikke følt dette har vært en hindring for utveksling av ideer rundt den foreslåtte løsningen. Takk for alle nyttige kommentarer og tilbakemeldinger.

Marianne Hove

Sammendrag

Europa-parlamentets og Rådet for Den Europeiske Unions direktiv av 15.mars 2006 ("Datalagringsdirektivet"), er utgangspunktet for dette arbeidets forslag til en teknisk og administrativ modell for innsamling og bruk av elektroniske kommunikasjonsdata. Datalagringsdirektivet krever at alle tilbydere av elektronisk kommunikasjon skal lagre trafikkdata som oppstår ved anvendelse av elektroniske kommunikasjonstjenester som fasttelefoni, mobiltelefoni, bredbåndstelefoni, internettaksess og e-post, og som kan identifisere lokasjon, bruker og utstyr som benyttes. Dataene skal lagres fra minst seks måneder og opptil to år. En innføring av direktivet betyr at en ny lagringsløsning, enten lokalt hos hver tilbyder, en sentral fellesløsning eller en mellomløsning må etableres. Dataene som lagres skal lagres separat fra data som brukes til fakturering av brukere av elektronisk kommunikasjon og samtrafikk. Nytteverdien kontra mulige negative effekter på personvernet en innføring av direktivet vil medføre, debatteres mellom politiet, politikere og interesseorganisasjoner. Dette arbeidet beskriver en lagringsløsning som tar utgangspunkt i standarder for håndtering av informasjonssikkerhet og rammeverk for lagring av trafikkdata samt dagens praksis for uthenting og bruk av elektroniske kommunikasjonsdata, og det settes opp en rekke krav og brukerscenarier for løsningen. Kravene er både administrative og tekniske, og brukerscenariene viser hvilke funksjoner den nye lagringsløsningen skal støtte. Det antas at dataene lagres kryptert, og nye mekanismer for søk i, uthenting og bruk av data må opprettes. Flere kryptoalgoritmer som støtter effektive søk i databaser, uten å avsløre søkestrengen eller resultatets innhold i klartekst har blitt vurdert. Dette arbeidets presenterte forslag

til en ny lagringsmodell ble basert på Hidden-Vector Encryption (HVE). HVE bruker offentlig nøkkel-kryptografi for aksesskontroll og baserer søk i en kryptert database på predikater. Den skjuler til en hver tid innhold i søkestreng og resultat, som kun kan dekrypteres ved besittelse og bruk av en skjult, privat nøkkel. Sikkerheten og kostnadene for den presenterte modellen har blitt vurdert, og modellen med bruk av HVE ble sammenlignet med en ukryptert databaseløsning. Løsningen ble også evaluert med tanke på misbruk av data gjennom innsideangrep.

Stikkord - Datalagringsdirektivet, ETSI-standard, Hidden-Vector Encryption, Predikater, ISO-standard, Personvern, Public Key Encryption with Conjunctive Keyword Search.

Innhold

Tabeller

Figurer

Forkortelser

CC	Common Criteria
CDR	Call Detail Record
DLD	Datalagringsdirektivet
DoS	Denial of Service
ETSI	European Telecommunications Standards Institute
IBE	Identity-Based Encryption
IEC	International Electrotechnical Commission
ISMS	Information Security Management Systems
ISO	International Standards Organisation
ISP	Internet Service Provider
HVE	Hidden Vector Encryption
MNO	Mobile Network Operator
MRQED	Multi-dimentional Range Query over Encrypted Data
MVNO	Mobile Virtual Network Operator
PT	Post- og Teletilsynet
PECKS	Public Key Encryption with Conjunctive Keyword Search
RSA	Rivest, Shamir, Adleman algoritmen
TR	Technical Report

Del I

Bakgrunnsinformasjon

Kapittel 1

Introduksjon

1.1 Formål med arbeidet

Dette arbeidet skal foreslå en teknisk og administrativ løsning som kan ivareta informasjonssikkerhet og personvern i beste måte, ved lagring av, søk i og uthenting av data, dersom EUs datalagringsdirektiv eller en lignende datalagringspraksis skulle bli vedtatt i Norge. For et best mulig vern om privatpersoner, deres kontaktnett og bevegelsesmønster må informasjonssikkerhetikkerhet være en av topprioritetene dersom direktivet innføres, hvilket betyr at sikkerheten både rundt datalagringen og uthentingen av data må være på topp. Den aktuelle situasjonen i samfunnet som danner bakgrunnen for dette arbeidets presentasjon av en ny lagringsløsning, er altså situasjonen der datalagringsdirektivet, eller lignende praksis, er vedtatt og innført i Norge, og tilbydere av elektronisk kommunikasjon er lovpålagt lagring av trafikkdata. Når det derfor senere refereres til den aktuelle situasjonsbeskrivelsen, eller det aktuelle scenariet, henvises det til en ny lovpålagt lagringsplikt, og all ny praksis som følger av denne.

I lagringsløsningen som presenteres i dette arbeidet er god sikkerhet prioritert fremfor lave kostnader. Dette for å imøtekomme enkeltpersoner eller grupperinger som stiller seg skeptisk til lagringen, samt for å øke tiltro til

både organiseringen og teknologien rundt datalagringen.

Det påstås ikke at å lagre data i henhold til direktivet er bedre enn å ikke lagre, eller omvendt. Dette arbeidet prøver ikke å si noe om fordelene eller ulempene veier mest ved en innføring av datalagringsdirektivet i Norge, ei heller å konstatere at direktivet vil bli innført. Formålet er kun å presentere et forslag til en administrativ og teknisk lagringsløsning. Løsningen skal kunne håndtere forespørsler om tilgang til dataene, kryptere dataene som lagres, søke i de krypterte dataene og dekrypte søkeresultatet, *dersom* direktivet eller lignende praksis skulle blitt innført i Norge nå, eller i fremtiden. Da formelt forslag til innføring av direktivet i Norge ennå ikke er lagt frem, er kommentarer og forslag til lagrings- og implementeringsløsning basert på direktivteksten, norske lovtekster og dokumentasjon på dagens praksis for uthenting og bruk av trafikkdata i etterforskningssaker.

1.2 Metode

For å besvare problemstillingen og formålet med dette arbeidet på en best mulig måte, ønskes arbeidet utformet som en tematisk, teoretisk studie. Brukereksempler og kravspesifikasjon til en løsning kan settes opp basert på et utvalg av bakgrunnsmateriale, og en løsningsmodell utformes basert på disse. Et sett med relevante og nødvendige artikler samles inn med formål om å sette opp en modell for en lagringsløsning ved eventuell innføring av datalagringsdirektivet i Norge. Selve direktivteksten og det tilhørende høringsnotatet fra Regjeringen bør studeres, og søk etter eksempler på praksis rundt kryptert datalagring, som kan være til hjelp under opprettelsen av en modell for en egen lagringsløsning, bør utføres. Flere alternative kryptoalgoritmer som kan passe til dette arbeidets formål om informasjonssikkerhet ved datalagring skal vurderes, og den best egnede til dagens aktuelle situasjon bør velges. Ulike brukseksempler (use cases) og en kravspesifikasjon basert på dette bakgrunnsmaterialet må så utformes, en teknisk modell for en lagringsløsning basert på disse modellene og kravene kan så presenteres. Lagringsløsningen skal være tilpasset direktivet, norsk lovgivning og praksis

for uthenting av data i dag, samt verne best mulig om privatpersoners rett til et privatliv.

Det ønskes ikke å utføre noen praktiske simuleringer av den presentere lagringsløsningen. På grunn av begrenset tid til dette arbeidet velges det heller å vurdere en mer teknisk, detaljert løsning på et kvalifisert teoretisk nivå, fremfor å utvikle og simulere en enklere, mindre detaljert modell. Praktiske simuleringer av modellen som skal presenteres kunne styrket de endelige konklusjonene, og gitt et bedre grunnlag for sammenligninger med andre måter å lage en løsning på. Ulike valg tatt i forbindelse med løsningen, samt sammenligninger med andre systemer, er basert på teori og eksperimenter redegjort for i andre akademiske artikler, samt kvalifiserte antakelser for hva som vil passe scenariet rundt datalagring best.

1.3 Organisering av arbeidet

For å besvare problemstillingen på en best mulig måte, er dette arbeidet organisert som følger. Kapittel 2 beskriver EUs Datalagringsdirektiv av 15.mars 2006 og generell informasjon om hva direktivet innebærer. Hvilke data som skal lagres, hvordan de er tenkt lagret og forskjellene direktivet medfører sammenlignet med dagens praksis for uthenting av elektroniske trafikkdata blir presentert. Nytteverdien av innføringen samt motstanden innføringen av direktivet møter, nevnes for å bedre begrunne de avgjørelsene som tas i forhold til den modellen som foreslås for lagringen. Kapittel 3 tar opp momenter med personvernet slik at man bedre kan vurdere hvilke hensyn som bør tas når en lagringsløsning foreslås. I kapittel 4 presenteres de standardene som er brukt som veiledning under arbeidet med å utvikle en modell for en ny lagringsløsning. En situasjonsmodell og noen brukseksempler (use cases) settes opp for å få en bedre oversikt over situasjonen vi står ovenfor, og for modellen som skal lages. Ved hjelp av standardene og de brukseksemplene som er laget, settes det i kapittel 5 opp administrative og tekniske krav til en eventuell lagringsløsning. Før egen modell presenteres, presenteres noen alternativer til kryptoalgoritmer som kan brukes i

lagringsløsningen i kapittel 6. Selve modellen for lagringsløsningen presenteres i kapittel 7. Løsningen baserer seg på å tilpasse den valgte kryptoalgoritmen til det spesifikke scenariet rundt datalagringsdirektivet. Kapittel 8 diskuterer fordeler og ulemper med den presenterte løsningen, og sammenligner den med de alternative kryptoalgoritmene som ble presentert i kapittel 6. Kapittel 9 konkluderer dette arbeidet.

Kapittel 2

Datalagringsdirektivet

Dette kapittelet omhandler EUs datalagringsdirektiv, hva direktivet går ut på, dets målsetninger, og de praktiske konsekvensene ved en eventuell innføring i Norge. Vi skal se på prosedyrer ved innsamling av data, sikkerhet og kostnader ved ulike lagringsmetoder, og forskjeller fra dagens lagringspraksis hos telefon- og internettilbydere, samt nevne noe om nytteverdien av en innføring av direktivet. Det er viktig å se på hva direktivet innebærer på ulike plan for å kunne sette opp et eventuelt forslag til en løsning ved dets innføring.

2.1 Generell informasjon

Europa-parlamentets og Rådet for Den Europeiske Unions direktiv 2006/24/EF av 15.mars 2006 [1], omhandler lagring av data som er generert eller behandlet i forbindelse med tilveiebringelse av offentlig tilgjengelige elektroniske kommunikasjonstjenester eller elektroniske kommunikasjonsnett. Dette direktivet gjenkjennes som *datalagringsdirektivet* (DLD). Data som skal lagres er trafikkdata, lokaliseringsdata og abonnement/brukerdata som oppstår ved anvendelse av elektroniske kommunikasjonstjenester som fasttelefoni, mobiltelefoni, bredbåndstelefoni, internettaksess og e-post. Dataene skal i henhold til direktivet *kun* utleveres til kompetente nasjonale myn-

digheter i særlige saker, og i overensstemmelse med nasjonal lovgivning under full overholdelse av de berørte personers grunnleggende rettigheter. Særlige saker sees på som "alvorlig kriminalitet", hvilket defineres av hvert enkelt medlemsland. Der tilbydere av elektroniske kommunikasjonstjenester før har lagret nødvendige data i en minimumsperiode for å fakturere kunder, skal det nå lagres større mengder data med flere detaljer over en lengre periode for å bekjempe kriminalitet.

Direktivet skulle i første omgang implementeres av alle EUs medlemsland innen 15.september 2007, men lagring av data fra e-post, internettkommunikasjon og internettelefon kunne utsettes til 15.mars 2009. Norge, som ikke er medlem av EU, forholder seg til de bestemmelser som gjelder under EØS-avtalen, hvor datalagringsdirektivet foreløpig ikke er innlemmet [2].

2.1.1 Definisjoner

I Datalagringsdirektivet defineres noen ord som kan være tvetydige. Disse definisjonene brukes også gjennom dette arbeidet ved omtale og diskusjon av direktivet og dets konsekvenser:

- Data: Trafikkdata, lokaliseringsdata og lignende data som er nødvendige for å identifisere abonnenten eller brukeren.
- Bruker: En hver juridisk eller fysisk person som anvender offentlig tilgjengelige elektroniske kommunikasjonstjenester til privat eller forretningsmessig bruk uten og nødvendigvis ha abonnement på tjenesten.
- Telefontjeneste: Oppkall (taleoppkalling, personsvar, konferanseoppkall eller dataoppkall), supplerende tjenester (viderestilling og omstilling) samt beskjed- og multimedietjenester (SMS-, EMS-, og MMS-tjenester).
- Brukeridentitet: En entydig identifikator som tildeles en person når vedkomne tegner abonnement eller lar seg registrere som bruker av en internettilgang eller internettkommunikasjonstjeneste.

- Celle-ID: Identiteten på en celle et mobilanrop kommer fra eller går til.
- Forgjeves oppkalling: Et telefonoppkall hvor det oppnås forbindelse, men som ikke besvares eller der nettverkssystemet har grepet inn.
- A-nummeret: Telefonnummeret til personen som originerer en samtale.
- B-nummeret: Telefonnummeret til den som mottar en samtale
- C-nummeret: Telefonnummeret til en som mottar en viderekoblet samtale.
- Tilbyder: tilbyr elektroniske kommunikasjonstjenester til kommersielt bruk for private kunder eller bedriftskunder.

Når disse ordene blir brukt i dette arbeidet, er det disse definisjonene det refereres til. Heretter, når det refereres til "direktivet", menes det aktuelle datalagringsdirektivet, Europa-parlamentets og Rådet for Den Europeiske Unions direktiv 2006/24/EF av 15.mars 2006 [1]. Når det refereres til "de lagrede dataene", refereres det til data lagret i henhold til direktivet, altså data separert fra de data som brukes av tilbydere til fakturering av kunder i dag.

2.1.2 Mål med direktivet

Målene med direktivet er å harmonisere medlemslandenes bestemmelser om pliktene som er pålagt tilbydere av elektroniske kommunikasjonstjenester vedrørende lagring av data de behandler, samt å sikre og kontrollere adgangen til disse dataene i forbindelse med etterforskning, avsløring eller rettsforfølgelse av grov kriminalitet, definert av de enkelte medlemsland [1]. Data som skal lagres inkluderer trafikkdata som er nødvendige for å identifisere og lokalisere abonnenten, men inkluderer ikke innholdet i selve kommunikasjonen. Målet er altså å bekjempe terrorisme og alvorlig kriminalitet gjennom lagring av trafikkdata nok til å identifisere og lokalisere en abonnent og det

anvendte kommunikasjonsutstyr (inkludert mobiltelefon og basestasjoner) på et visst tidspunkt.

2.2 Innsamling og lagring av data

Deler av informasjonen som kreves lagret av direktivet, lagres også allerede i dag [3]. Noe data som ikke lagres i dag genereres likevel hos tilbyderne, slik at informasjonen er mulig å lagre dersom lagring skulle bli lovpålagt. Dataene som skal lagres er informasjon nok til at abonnenter kan identifiseres og lokaliseres, men innholdet i selve kommunikasjonen skal ikke lagres. Sikkerheten rundt den nye lagringsløsningen vil bli et hovedfokus for folk som er opptatte av personvernet, mens kostnadene vil bli et hovedfokus for tilbyderne. Dette er fordi kostnadene ved en ny løsning kan få store betydninger for konkurransen i markedet. Kostnadene ved ulike teknologier og ulike typer lagringsløsninger kan påvirke tilbydernes prioriteringer dersom dette valget blir opp til dem. Valg av løsning vil også ha noe å si for hvordan prosessen rundt å hente ut dataene blir for politiet.

2.2.1 Hva skal lagres?

Dersom direktivet innføres i Norge, skal det lagres mer data enn det allerede gjør med dagens praksis. Direktivets artikkel 5 presenterer hvilke opplysninger som skal lagres (Tabell 2.1).

I følge direktivet artikkel 6, skal dataene lagres i minst seks måneder og høyst i to år fra datoen for kommunikasjonen. Hvert medlemsland bestemmer selv hvor lenge innenfor dette spekteret. Sammenlignet med hva som lagres hos tilbydere av telekommunikasjon og internettilgang i dag, skal det lagres *flere* opplysninger, og de skal lagres *lenger* enn ved dagens praksis [2]. I dag lagres A-nummeret (nummeret til den som initierer en samtale) samt navn og adresse på abonnent eller registrert bruker for faktureringseller kommunikasjonsformål. Dette betyr at abonnentens identitet må lagres, men ikke nødvendig vis hvilke telefoner eller hvilket utstyr som brukes

Tabell 2.1: Data som skal lagres i henhold til Datalagringsdirektivet.

	Fast- og Mobiltelefoni.	*Internett, E-post og VoIP.*
Data nødvendig for å spore/identifisere kilden i kommunikasjonen.	*A-nummeret* *Navn og adresse på registrert bruker*	*Tildelt brukeridentitet.* *Navn og adresse på abonnent eller registrert bruker, samt IP-adresse.* *brukeridentitet eller telefonnummer tildelt på kommunikasjonstidspunkt.*
Data for å fastslå kommunikasjonens bestemmelsessted.	*B-nummeret.* *Navn og adresse på abonnenten eller den registrerte brukeren.*	*Brukeridentitet og telefonnummer.* *Navn og adresse på abonnenten eller den registrerte brukeren.*
Data for å fastslå kommunikasjonens dato, klokkeslett og varighet.	*Dato og klokkeslett for kommunikasjonens begynnelse og sluttidspunkt.*	*Dato og klokkeslett for inn- og utlogging av internettjenester basert på en tidssone og dynamisk eller statisk IP-adresse, samt brukeridentitet på abonnent eller registrert bruker.* *Dato og klokkeslett for inn- og utlogging av e-posttjeneste basert på en bestemt tidssone.*
Data nødvendig for å identifisere kommunikasjonstypen.	*Den anvendte telefontjeneste.*	*Den anvendte internettjeneste.*
Data nødvendig for å identifisere brukerens kommunikasjonsutstyr, eller det som fremstår som brukerens utstyr.	*Ved forhåndsbetalte anonyme tjenester, dato og tidspunkt for første aktivering og lokaliseringskode (celle-ID) hvorfra aktiveringen ble foretatt.* *A-nummeret og B-nummeret.* *A-abonnentens IMSI nummer (Internasjonal Mobilabonnements Identitet.)* *A-abonnentens IMEI nummer (Internasjonal Mobilutstyrs Identitet).* *B-abonnentens IMSI.* *B-abonnentens IMEI.*	*For internett og e-post: A-nummeret mtp dial-up-tilgang, samt den digitale abonnementslinjen (DSL) eller annet endepunkt for kommunikasjonens opphavsperson.* *For VoIP: A-nummeret og B-nummeret.*

Data nødvendig for å foreta lokalisering av mobilt utstyr	*Celle-ID ved kommunikasjonens begynnelse. Data som med henvisning til celle-ID viser cellens geografiske lokalisering i den perioden hvor det lagres kommunikasjonsdata.*	

når abonnent/bruker skaper trafikk i nettet. Når det gjelder internettilbydere kan de være pliktige til å lagre identiteten til abonnenter på IP-telefoni, men trafikkdata skal ikke lagres. Både for operatører innenfor telekommunikasjon og internettilgang er lagring av dataene regulert av personopplysningsloven [4] (se kapittel 3) og Ekomloven [5]. Ekomloven §2-7 annet ledd oppstiller en sletteplikt for trafikkdata når dataene er overflødige med tanke på fakturerings- og kommunikasjonsformål: "Trafikkdata skal slettes eller anonymiseres så snart de ikke lenger er nødvendig for kommunikasjons- eller faktureringsformål, med mindre annet er bestemt i eller i medhold av lov. Annen behandling av trafikkdata krever samtykke fra bruker." Maksimal lagringstid er fem måneder etter kvartalsvis fakturering er registrert, og tre måneder etter månedlig fakturering er registrert. Dersom fakturaer ikke betales eller en rettslig tvist oppstår, kan dataene likevel lagres inntil kravet er rettlig oppgjort. Også i datalagringsdirektivet er det inkludert en sletteplikt. I direktivets artikkel 7, avsnitt d) står det at dataene skal tilintetgjøres ved utløp av lagringstiden, med mindre dataene har vært gitt tilgang til og gjemt bort.

Direktivets artikkel 5 nr. 2, konstaterer at data som avslører innholdet i kommunikasjonen *ikke* skal lagres i følge direktivet. For en mest mulig vennlig lagring sett fra personvernets side vil man også unngå lagring av sensitive data. Sensitiv informasjon er i følge personopplysningsloven §2-8 informasjon som inneholder opplysninger om rasemessig eller etnisk bakgrunn, politisk eller religiøs oppfatning; opplysninger om en person har vært misstenkt, siktet, tiltalt eller dømt for en straffbar handling; og opplysninger som beskriver helseforhold, seksuelle forhold eller medlemskap i

fagforeninger [4].

2.2.2 Sikkerhet ved lagring av data

I følge datalagringsdirektivet artikkel 7 skal data som lagres ha samme kvalitet og være omfattet av samme sikkerhet og beskyttelse som de data som genereres i det elektroniske kommunikasjonsnettet. Dataene skal være omfattet av nødvendige tekniske og organisatoriske foranstaltninger slik at de er beskyttet mot utilsiktet eller ulovlig tilintetgjørelse eller utilsiktet tap mot degradert, uautorisert eller ulovlig lagring, behandling, adgang eller utbredelse. Det skal sikres at kun særlig autoriserte personer får adgang til dataene. Datalagringsdirektivet sier altså ingenting om krav til maksimum nedetid eller minimum oppetid eller responstider for uthenting av data. Direktivet gir heller ingen presise retningslinjer for hvor sikkert dataene skal lagres, kun at de skal være like godt sikret som samme type data allerede er.

Politiet skal ha tilgang til data om aktuelle abonnenter uten en beslutning fra retten, men kun tilgang til trafikkdata, inkludert varigheter på samtaler, tidsrom, lokasjon og kommuniserende abonnenter og utstyr; dersom de har en rettslig beslutning om dette. Disse ulikt begrensede tilgangene, samt at informasjonen som er lagret bør krypteres for et enda bedre vern om personers privatliv, kan gjøre sikker lagring utfordrende.

Et annet aspekt ved sikkerheten under lagring er lokasjoner det blir lagret på, og hvor mye data som lagres på hver lokasjon. I regjeringens høringsnotat går de bort fra å foreslå at en sentral lagringsløsning skal etableres, selv om dette ville være billigste alternativ i følge Teleplans undersøkelse fra 2006 [6]. I en slik løsning ville alle tilbydere ha overført sine data i et standardisert format til en sentral database en gang per døgn. Et annet alternativ er lagring kun hos tilbyder, der det hos Teleplan forutsettes at det opprettes egne lagringsløsninger for data lagret i henhold til direktivet, adskilt fra de data tilbyderne allerede lagrer og bruker til blant annet fakturering. En presentert mellomløsning innebærer at de små tilbyderne kan gå sammen om en sentral løsning for å dele kompetanse og ressurser, mens de som har

råd til det bygger på sine egne, nåværende systemer for lagring. Forskjellen ved sikkerheten til de ulike lagringsløsningene ligger ved hvem som han ansvar for lagringen. I en sentralt drevet løsning, må det nødvendigvis være en tredjepart som er ansvarlig for lagringen, inkludert sikkerhet og vedlikehold. Dersom tilbyderne lagrer hos seg selv, eller eventuelt går sammen om en felles løsning, vil de kun måtte forholde seg til minimumskravene hva sikkerhet angår, og implementere ulike løsninger som tilfredsstiller kravene til sikkerhet rundt datalagring.

2.2.3 Kostnader ved lagring av data

Denne seksjonen tar for seg kostnadene, som i utgangspunktet påføres tilbyderne, som oppstår dersom en ny datalagringspraksis innføres i Norge. Lagringsalternativene det er mulighet for å velge mellom resulterer i ulike kostnader, og hvordan fokus på teknologi og informasjonssikkerheten prioriteres vil også påvirke kostnadsnivået. Valg av teknologi vil påvirke tiden det tar å utvikle løsningen, som igjen også vil påvirke kostnadene.

Ulike lagringsløsninger vil gi ulike kostnader for ulike aktører. De lagringsløsningene som er presentert av Teleplan i 2006 og i Regjeringens høringsnotat er en sentral lagringsløsning, en distribuert lagringsløsning hos hver tilbyder eller en mellomløsning (se også seksjon 2.2.2). I følge Teleplans utredning er en sentral lagringsløsning det beste økonomiske alternativet for tilbyderne, og mellomløsningen med muligheter for sentral lagring og lokal lagring blir det dyreste alternativet. Uansett vil de ulike løsningene implementeres forskjellig og utviklingstiden for etablering av løsningen vil være forskjellig for de ulike alternativene. Dette er det ikke tatt høyde for i Teleplans utredning. Ekstra separat lagring for direktivets data i tillegg til den lagringen som gjøres hos tilbydere i dag, vil gi økte kostnader for tilbyderne. Ikke bare utviklingen av løsningen vil koste penger, men vedlikehold og opplæring av ansatte til å drive løsningen vil også kreve tid og penger. Utviklingen og etableringen av løsningen som velges vil være av større betydning hva kostnader angår enn lagringstid og -volum vil være [6]. Samtidig kan en ny lagringspraksis gi reduserte kostnader andre steder. De

kostnadene politiet har i dag gjennom å skaffe dokumentasjon på organisert kriminalitet, som overvåkning, kan reduseres ved innhenting og analyse av trafikkdata.

2.3 Forskjeller fra dagens praksis

Hovedforskjellene på dagens praksis og det som vil bli praksis dersom direktivet innføres, er hvor detaljert innhold dataene som lagres skal ha, hvor lenge de skal lagres og formålet med lagringen. Tilganger til eller uthenting av data skal i hovedsak reguleres av hvert enkelt medlemsland i henhold til nasjonal lovgivning, og en innføring av direktivet vil ikke innebære store endringer fra dagens praksis vedrørende tilgang eller uthenting av elektroniske kommunikasjonsdata.

2.3.1 Dagens lagringspraksis

Informasjonen som lagres i dag er informasjon som er nødvendig for fakturering av sluttbrukere, beregning av samtrafikkpartnere og kapasitetsovervåkning. I hvilket format dataene lagres og hvor lenge, er avhengig av formålet med lagringen og hvilke systemer tilbydere av kommunikasjonstjenesten bruker. Formatene på lagrede data kan også variere fra tilbyder til tilbyder, da tilbydernes behov og operative prosesser må tilpasses. Noen lagrer såkalte Call Detail Record (CDR) direkte, andre konverterer til andre formater for tilpassede lagringsmetoder. CDR inneholder identiteten til kildene for kommunikasjonen og identiteten til destinasjonen/endepunktet, varigheten av en samtale, hva som skal faktureres for hver samtale, total brukstid i faktureringsperioden og totale kostnader.

Teleplan har allerede utført en undersøkelse på hvilke data som lagres i dag blant aktuelle tilbydere av fasttelefoni, mobiltelefoni, bredbåndstelefoni, bredbånd og internettaksess m/e-post [3]. Undersøkelsene gjort av Teleplan viser at det for fasttelefoni allerede lagres eller genereres A-nummer, kundeidentitet og B-nummer. C-nummer lagres hos B-nummerets tilbyder, og noen tilbydere lagrer start og sluttid for en samtale, mens andre lagrer

kun varigheten. For mobiltelefoni lagres eller genereres A-nummer, kundeidentitet og B-nummer av alle, mens C-nummer lagres hos B-nummerets tilbyder. Samtlige spurte tilbydere lagrer IMSI, men under halvparten lagrer IMEI. Under en tredel lagrer begge, og hos de som ikke lagrer verken IMEI eller IMSI, genereres de heller ikke. Kun nettverksoperatører (Telenor og NetCom) lagrer data om celle-ID. Hos tilbydere av IP-telefoni lagres eller genereres A-nummer, kundeidentitet og B-nummer. C-nummer lagres hos B-nummerets tilbyder, og noen tilbydere lagrer varighet av samtale istedenfor start- og sluttidspunkt. Halvparten av tilbyderne lagrer identifikator for aksesslinje. For Internet Service Provider's (ISP), tilbydere av internettaksess, lagres data i mindre grad enn for fast- og mobiltelefoni. Litt over halvparten av de spurte tilbyderne lagrer IP-adresser og kundeidentiteter, og utover dette er det få av tilbyderne som genererer disse dataene. For ISPenes e-posttjenester lagres data i enda mindre grad enn for internettaksess. Under halvparten av de spurte lagrer nødvendige data, og litt over halvparten oppgir at de lagrer eller genererer disse dataene. Når det kommer til andre tilbyderes e-posttjenester, som Gmail, Hotmail, og Yahoo-mail, er det ingen som lagrer data i Norge. Mest sannsynlig lagrer eller genererer de data på fysiske lokasjoner utenfor Norge, men de omfattes da ikke av datalagringsdirektivet.

Hierarkiet av tjenestetilbydere i Norge av fast- eller mobiltelefon har også innvirkning på hva som lagres i dag og hvordan. Kun Telenor har eget nett for fasttelefon i Norge, mens både Telenor og NetCom har nett for mobiltelefoni og er såkalte Mobile Network Operators (MNO), nettverksoperatører. Dette betyr at alle andre tilbydere av mobile kommunikasjonstjenester må leie nett av enten Telenor eller NetCom, for så å selge tjenester til sine kunder. Slike tilbydere kalles ofte Mobile Virtual Network Operator (MVNO), og er bare virtuelle nettverksoperatører. Når tilbyderne uten eget nett selger sine tjenester vil de selv besitte informasjon om sluttbruker og abonnementsdata som navn og adresse, men aktiviteten på nettet er det netteier/nettilbyder som besitter. Kommunikasjonsdata kan dermed eksistere på flere lokasjoner hos flere tilbydere samtidig i enkelte perioder. I dages situasjon er det dermed ikke sikkert alle tilbydere alene sitter på den informasjonen datala-

gringsdirektivet kreves lagret. Men på grunn av samtrafikkavregning vil all informasjon være lagret hos Telenor og NetCom i tillegg til hos de som leier nett av disse. I følge regjeringens høringsnotat forholder politiet seg derfor kun til Telenor og NetCom for uthenting av kommunikasjonsdata, da disse er de med eget fast- og eller mobilnett i Norge.

2.3.2 Tilgang til og uthenting av lagrede data i dag

I følge straffeprosessloven §215a kan påtalemyndigheten gi pålegg om sikring av elektronisk lagrede trafikkdata som antas å ha betydning som bevis i en etterforskning [7]. Det må være grunn til å tro at det er begått en straffbar handling. Sikring av data gir ikke automatisk rett for påtalemyndighet til å uthente elektronisk trafikkdata. Retten kan pålegge besitteren av *ting som antas å ha betydning som bevis* å utlevere disse. Trafikkdata inngår under *ting*. Påtalemyndigheten kan for øvrig gi ordre om utlevering uten rettslig kjennelse dersom det er fare for at etterforskningen vil lide. Denne beslutningen skal snarest forelegges retten for godkjennelse. I tillegg må Post- og Teletilsynet (PT) forespørres om fritak fra taushetsplikten i hvert enkelt tilfelle for utlevering av historiske kommunikasjonsdata, da tilbydere bryter taushetsplikten sin ved utlevering av data. Fritak av denne taushetsplikten, som er regulert av Ekomloven §2-9, kan bare nektes dersom åpenbaringen vil kunne utsette staten eller allmenne interesser for skade, eller virke urimelig for den som har krav på hemmelighold [5]. Dersom en person ikke er mistenkt eller siktet for straffbare forhold, vil det derfor ligge restriksjoner hos PT fra å gi fritak fra taushetsplikten. Hvor viktige elektroniske spor er for etterforskningen er også med i denne vurderingen. Kun dersom PT gir fritak kan politiet sende begjæring om utelvering av data til retten. Legg merke til at det i dag er forskjell på utlevering av elektronisk trafikkdata og utlevering av opplysninger om abonnenten. Politiet kan henvende seg direkte til tilbyder av elektroniske kommunikasjonstjenester for å uthente opplysninger om abonnenter, uten å innhente samtykke fra Post- og Teletilsynet jf. Ekomloven §2-9, tredje ledd. Politiet trenger ikke en rettslig beslutning for å hente ut informasjon om abonnenter fra teleoperatører eller ISPer, men de trenger en rettslig beslutning for å hente ut lagret trafikkdata.

Oppsummert vil dette si at kommunikasjonsdata som lagres og genereres i dag er litt forskjellig hos de ulike tilbyderne, både på grunn av ulike systemer som brukes og ulik tilgang til og nytteverdi av dataene for eksempel til faktureringsformål. Hierarkiet av tilbydere gjør at data også kan være lagret flere steder samtidig. Tilbydere som lagrer kommunikasjonsdataene i dag forholder seg til, og har lenge gjort det, Ekomloven [5] og personopplysningsloven [4], og forventes derfor å ha noe administrativ erfaring vedrørende å lagre og behandle data om enkeltpersoner (abonnenter) i henhold til norsk lovgivning.

2.4 Nytteverdien ved innføring av direktivet i Norge

En av grunnene til at EU vil innføre direktivet er at de mener en innføring av direktivet i meldemslandene vil få stor nytteverdi, spesielt innenfor kriminalitetsbekjempelse. Formålene med direktivet er å sikre lik tilgang til informasjon om elektroniske kommunikasjonstjenester og abonnenter i alle medlemsland, slik at etterforskning av grov kriminalitet kan bli bedre og grundigere. (Hva som regnes som grov kriminalitet skal for øvrig defineres av de enkelte medlemsland). Nytteverdien av direktivets innføring vil altså ligge i bedre etterforskningsmuligheter og bedre internasjonalt politisamarbeid i saker som omhandler grov kriminalitet. I likhet med resten av samfunnet brukes elektroniske kommunikasjonstjenester i stadig økende grad. I kampen mot alvorlig kriminalitet kan tilgang til mer detaljert data om slik kommunikasjon kunne kartlegge kriminelle nettverk og muligens forsterke argumentene for eller mot i en straffesak. Dersom alle EUs medlemsland innfører direktivet, hvilket de i utgangspunktet skal, vil alle "kompetente nasjonale myndigheter" (som i mange tilfeller vil være politiet) i medlemslandene ha tilgang til de samme typer data i minimum 6 måneder. Svein Willassen har på oppdrag fra Datatilsynet skrevet en rapport om Datalagringsdirektivets verdi i etterforskning og risikofaktorer for personvernet [8]. Vedrørende datalagringsdirektivets innføringsverdi i etterforskning kommer Willassen frem til at logging av internetttilgang, altså hvem som tilhør-

er hvilken IP-adresse på et gitt tidspunkt, vil være viktig i etterforskningssammenheng. Også lagring av trafikkdata fra mobiltelefoner, fastnett og IP-telefoni ansees som verdifulle. Nytteverdien av logger over hvem som sender e-post til hvem ansees som lite nyttig.

2.4.1 Bekjempe alvorlig kriminalitet

De siste tiårs utvikling innenfor teknologi har gjort at metodene for kommunikasjon og spredning av informasjon har endret seg både for dagligdags bruk og også for kriminelt bruk. Nye metoder brukt til terror og andre former for kriminalitet krever også nye metoder ved etterforskning og bekjempelse av dette. Politiet og påtalemyndigheter trenger derfor tilgang til nye teknologiske verktøy, men nye metoder bør være klart regulert etter demokratiske regler. Spørsmålet er om det omfanget av data direktivet krever lagret gir så store forbedringer innenfor etterforskning av alvorlig og organisert kriminalitet og oppklaring av saker, at det veier opp for ulempene lagringen fører med seg.

Det er flere ulike datatyper som skal lagres i henhold til direktivet (se seksjon 2.2.1), og de kan ha ulik betydning i forhold til etterforskning. Man kan skille mellom telefonilogger, e-postlogger, og internettlogger [9]. Telefoniloggene inkluderer abonnementsinformasjon, lokasjon og varighet under samtaler fra fast- og mobiltelefoner; e-postloggene inkluderer navn/bruker-ID fra e-poster som blir sendt; og internettloggene inkluderer abonnementsinformasjon relatert til internettaksess, for eksempel IP-adresser. Telefoniloggene brukes allerede som bevis i saker under etterforskning, og uthenting av dataene reguleres av straffeprosessloven [7] (se seksjon 2.3). Det blir her gjort forskjell på om det er snakk om å hente ut informasjon om abonnenten eller trafikkdata til en gitt abonnent. Det gjøres også forskjell på telefonilogger som allerede er lagret (sikring av bevis) og overvåkning til senere bevisbruk (kommunikasjonskontroll). Statistikk for politiets forespørsel etter trafikkdata er ikke klarlagt. I følge Telenor, på forespørsel fra Datatilsynet, er det deres politisvarsenters generelle inntrykk at trafikketterspørsel fra politiets side gjelder tidsperioder knyttet til de siste 3 måneder [2]. Te-

lenor påpeker også at politiet ikke fremmer spørsmål om data eldre enn 3 måneder, fordi de vet Telenor ikke har slike data. Det er derfor vanskelig å si noe om i hvilken grad politiet ville fått bruk for data eldre enn 3 måneder i en etterforskning. E-postlogger brukes sjelden som bevis i dag siden hele e-poster med innhold som regel lagres på maskiner hos både sender og mottaker, og kan derfor finnes og beslaglegges i etterforskningssammenheng. Når det gjelder IP-adresselogger, der det skal lagres navn og adresse på den registrerte bruker og tildelt(e) IP-adresse(r), sier dette lite om lokasjon og selve bruken av internett. I motsetning til før, da man måtte ringe opp via modem for å få internettforbindelser, er man i dag tilkoblet internett kontinuerlig og betaler gjerne månedspris for dette til en ISP. Dette er også en av grunnene til at ISPer ikke har samme behov for logging som teleoperatørene, som fakturerer basert på bruk. Men hvem som tilhører hvilken IP-adresse kan være viktig informasjon i etterforskningssammenheng. Mengden virus, angrep gjennom tjenestenekt, spredning av barneporno og andre overtredelser av loven øker, og uten å vite hvem som står bak hvilke IP-adresser, kan vi til en viss grad oppleve at håndhevelse av loven kan være forgjeves [9]. Lengre lagring av brukeridentiteter som er knyttet til IP-adresser kan derfor være nyttig i fremtidig etterforskning. Fordi fakturering i hovedsak skjer på månedsbasis (altså at bruken ikke logges i samme grad om hos teleoperatørene), og lagret info sier lite om abonnentens lokasjon/bevegelse (dersom man har statisk IP-adresse vil man kunne bevege seg uten at geografisk lokasjon av IP-adresse endres) har kanskje lagring av IP-adresselogger en annen effekt på personvern enn telefoniloggene også.

2.4.2 Bedre internasjonalt politisamarbeid

Lik tilgang på informasjon i alle medlemsland er et av målene med datalagringsdirektivet. Direktivet skal harmonisere medlemslandenes bestemmelser om pliktene de har til lagring av elektroniske kommunikasjonsdata. I følge Politiet har det vært flere tilfeller hvor de har mottatt informasjon, for eksempel IP-adresser, fra internasjonale kolleger, men hvor de ikke har klart å finne tilhørende brukere på grunn av for kort lagringstid hos norske ISPer [10]. Norsk politi hadde da ikke lik tilgang på informasjon som sine

internasjonale kolleger, og det ble i nevnte tilfelle tatt flere brukere i andre land. Ved innføring av direktivet vil alle medlemsland ha tilgang på en minimumsmengde data i minst 6 måneder. Noen land kan bestemme seg for å innføre og lagre mer informasjon enn det som kreves i direktivet, og noen land vil lagre dataene i 6 måneder, andre i 2 år. Så det vil med andre ord fortsatt være ulike mengder data og lagringstider i de ulike medlemslandene. Likevel vil en innføring av direktivet i Norge medføre at internasjonalt politi vet hva de kan forvente av norsk politi og deres mulige informasjonstilgang i etterforskning eller straffesaker, samt at politiet i Norge kan forvente det samme av sine internasjonale kolleger.

2.5 Direktivet møter motstand

Direktivet har siden det ble vedtatt den 15. mars 2006 møtt kraftig motstand over hele Europa. Både på grunn av personvernhensyn og tekniske sikkerhetsaspekter. Den tverrpolitiske uavhengige organisasjonen *Stopp Datalagrinsdirektivet* har blant annet i sin uttalelse påpekt mange av de negative sidene ved en innføring av direktivet i Norge, og fraråder regjeringen å innføre det [11]. Sikkerheten ved lagring og uthenting av data fryktes nedprioritert til et minimum når kostnadene for bedriftene øker pga lagring av større mengder data over lengre tid enn hva som praktiseres i dag.

2.5.1 Datalagringsdirektivet og personvern

Lengre lagring av mer detaljert kommunikasjonsdata enn hva som praktiseres i dag gjør at det stilles nye spørsmål vedrørende et svekket personvern. Registrering over en lengre tidsperiode av vår lokasjon og hvem våre samtaler er med og hvor lenge, ville kunne gitt et godt bilde av vårt bevegelsesmønster og kontaktnett dersom hele vår kommunikasjonshistorie over minst 6 måneder ble kjent. Noen vil kanskje kjenne integriteten sin krenket bare ved å vite at noen faktisk besitter slik informasjon, og frykten øker ytterligere med tanke på at denne informasjonen faktisk kan misbrukes. Denne konsekvensen av direktivets eventuelle innføring påvirker personvernet (Kapittel 3), og derfor må både personopplysningsloven

og Den Europeiske Menneskerettighetskonvensjonen (EMK) respekteres. I EMK artikkel 8 lyder det:

1. Enhver har rett til respekt for sitt privatliv og familieliv, sitt hjem og sin korrespondanse.

2. Det skal ikke skje noen inngrep av offentlig myndighet i utøvelsen av denne rettighet unntatt når dette er i samsvar med loven og er nødvendig i et demokratisk samfunn av hensyn til den nasjonale sikkerhet, offentlige trygghet eller landets økonomiske velferd, for å forebygge uorden eller kriminalitet, for å beskytte helse og moral, eller for å beskytte andres rettigheter og friheter.

Her er det tydelig at Datalagringsdirektivet vil omfattes av EMK artikkel 8 nr. 1, men spørsmålet er om lagringen av dataene i det tidsrommet direktivet omhandler likevel kan rettferdiggjøres av EMK artikkel 8 nr. 2.

I følge punkt (11) i innledningen av direktivet er det i undersøkelser i EUs medlemsland påvist at slik generert elektronisk trafikk- og lokaliseringsdata har stor betydning i etterforskning, avsløring og rettsforfølgelse av straffbare handlinger. Derfor mener Europa-parlamentet og Rådet for Den Europeiske Union at det er nødvendig å sikre at slike elektroniske kommunikasjonsdata er tilgjengelige for håndhevelse av loven for en bestemt periode, en periode fastsatt i direktivet. Nettopp fordi lagring av elektroniske kommunikasjonsdata har vist seg såpass nyttig i etterforskning, fastslår direktivet at offentlige myndigheter kan gripe inn i retten et hvert menneske har til privatliv og familieliv, fordi det er nødvendig med hensyn til nasjonal sikkerhet, offentlig trygghet, forebyggelse av uro og forbrytelser, eller for å beskytte andres rettigheter og friheter. Man kan også skille mellom effekten lagring av ulike typer data har på personvernet [8]. Logg over hvem som tilhører hvilken IP-adresse kan ansees som å påvirke personvernet i liten grad, da mange er kontinuerlig pålogget gjennom bredbånd, og av- og pålogging ikke nødvendig vis sier noe om når man selv logger av eller på. I mange husstander i dag står ruteren alltid koblet opp mot internett, men frakobles kanskje kun ved strømbrudd eller andre eksterne hendelser

[8]. Dermed sier ikke IP-adressene noe særlig om selve bruken av internett. Lagring av trafikkdata fra mobiltelefon, fastnett og IP-telefoni, samt logging av hvem som har sendt e-post til hvem, kan derimot ansees som å påvirke personvernet i større grad.

2.5.2 Sikkerhet ved lagring og uthenting nedprioriteres

De mange over 200 tjenestetilbyderne på offentlig ekomnett og offentlige teletjenester vi har i Norge i dag [12], er av svært ulik størrelse, med et ulikt antall tjenester og kunder. For de mindre tilbyderne kan de økende kostnadene ved lagring i henhold til direktivet bli konkurransehemmende. Spørsmålet er da om sikkerheten vil bli prioritert, når de ulike tilbyderne havner i pressede økonomiske situasjoner. Ikke alle aktører er like seriøse og det kan også bli vanskelig å kontrollere sikkerheten hos hver enkelt tilbyder. Sikkerheten rundt skjermingsverdig informasjon reguleres av loven om forebyggende sikkerhetstjeneste (sikkerhetsloven) [13]. Artikkel 4 omfatter informasjonssikkerhet. Dersom de lagrede dataene regnes som skjermingsverdig og sikkerhetsgraderes blir enhver som får tilgang til sikkerhetsgradert informasjon som ledd i arbeid, oppdrag eller verv for en virksomhet, pliktig i å hindre at uvedkommende får kjennskap til informasjonen. For å hindre at uvedkommende får tilgang, kan de ulike tilbyderne benytte seg av kryptosystemer. Disse kryptosystemene må være godkjent av Nasjonal sikkerhetsmyndighet. I følge Ekomloven §2-7 skal tilbydere av elektroniske kommunikasjonstjenester "...gjennomføre nødvendige sikkerhetstiltak til vern av kommunikasjon i egne elektroniske kommunikasjonsnett og -tjenester. Ved særlig risiko for brudd på sikkerheten skal tilbyder informere abonnenten om risikoen." Selv om informasjonssikkerhet for skjermingsverdig informasjon er lovregulert, kan de spesifikke systemkravene for lagring av og tilgang til disse opplysningene senkes til et minimum. Det kan antas at svært få tilbydere kan tenke seg å bruke mer penger enn nødvendig på en eventuell tvungen lagring av data, og vil derfor ikke bruke mer på et enda sikrere system enn nødvendig. Dette er bekymringsverdig med tanke på den verdien den lagrede informasjonen kan ha for uvedkommende.

Kapittel 3

Personvern

Personvern og personopplysningsvern i sammenheng med lagring av kommunikasjonsdata betyr å verne om personers integritet og sikring av at persondata, selvstendighet, uavhengighet og ukrenkelighet beskyttes. Hver enkelt person er selvfølgelig interessert i å beholde sitt privatliv, men også å sikre at de dataene som først lagres er korrekte og relevante til formålet (nå: avregning for samtrafikk, og fakturering av kunder). Artikkel 29-gruppen er en rådgivende arbeidsgruppe som er nedsatt ved Europa-Parlamentets og Rådets direktiv 95/46/EF av 24. oktober 1995 [14], dannet for å beskytte fysiske personer i forbindelse med behandling av personopplysninger. Denne gruppen har også utarbeidet en uttalelse vedrørende datalagringsdirektivet [15]. Personvern er deres fokus for uttalelsen, og de mener behov for utvidet lagring og tilgang til trafikkdata i det omfang som er foreslått av direktivet må begrunnes bedre. Også konsekvensene dersom direktivet *ikke* innføres vil de gjerne ha utredet. Artikkel 29-gruppen utelukker ikke at trafikkdata kan være viktige for politiet ved etterforskning og oppklaring av saker som inkluderer kriminelle handlinger, men de er fortsatt kritiske til beslutningsgrunnlaget for direktivet. De mener en innføring av direktivet setter både personvern og ytringsfrihet på prøve og at "...vissheten av at noen kan lete seg fram til dine kontakter og dine bevegelser, både i det virkelige rommet og på Internett, kan være nok til å hemme borgere i utøvelsen av sine fri-

heter til å samles, til å ytre seg og til å søke opplysninger." [15]. Muligheten for å lete seg fram i kontaktnett og bevegelser gjennom lagring av trafikkdata bør derfor begrenses så mye som mulig i form av at dataene som lagres krypteres, og at tilgangsmetodene er begrendsede og kontrollerte samt forebeholdt svært få brukere. For å verne best mulig om de dataene som skal lagres dersom datalagringsdirektivet eller lignende praksis skulle bli innført i Norge, er det viktig at personvern, gjennom god informasjonssikkerhet, blir prioritert i den eventuelle lagringsløsningen.

Om direktivet lovlig kan innføres ved lov med tanke på motstridende lover vedrørende personvern, som EMK (2.5.1) diskuteres ikke videre her. Men det noteres at situasjonen rundt motstridende lover er omstridt.

Del II

Forslag til løsning

Kapittel 4

Modellering av aktuell situasjon

Dersom datalagringsdirektivet blir innført i Norge, eller dersom et lignende direktiv med krav om lagring skulle bli innført senere, er det ønskelig med en implementeringsløsning som støtter norsk lovgivning med tanke på utlevering av og tilgang til dataene, samt at dataene er sikret best mulig med tanke på personvernet. Det er også ønskelig for bedre tiltro til implementeringsløsningen at valg av administrative og tekniske løsninger er standardiserte eller sertifiserbare slik at løsningen lettere kan evalueres og verifiseres av en nøytral tredjepart. I dette kapittelet presenteres generelle og noen mer spesifikke scenarioer og modeller for hvordan løsningen skal kunne brukes. Det settes opp en modell for hele situasjonen rundt lagringen sett under ett, og deretter flere mindre brukerscenarioer, eller use cases, som representerer ulike bruksområder og handlinger som lagringsløsningen skal kunne håndtere.

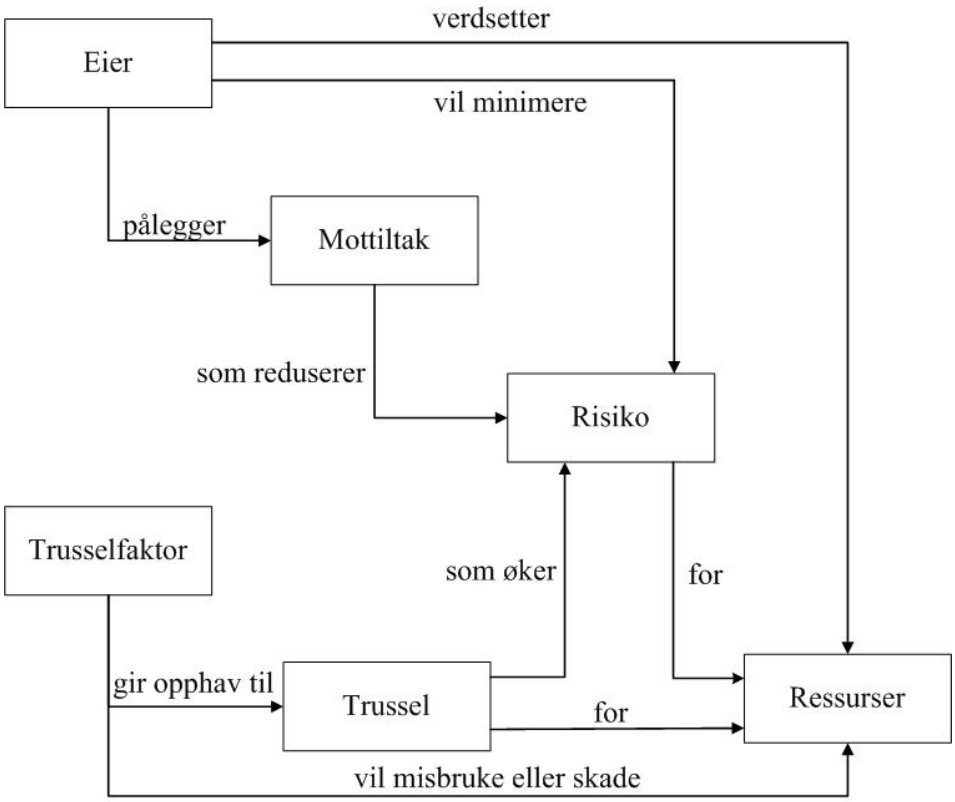
4.1 Standarder

Standarder fra International Standards Organization (ISO) og International Electrotechnical Commission (IEC), og European Telecommunications

Standards Institute (ETSI) har blitt brukt som veiledning i utredningen av brukseksempler, administrative og tekniske krav, og selve lagringsløsningen. Fra ISO/IEC har standard ISO/IEC 15408, Common Criteria (CC) for Information Technology Security Evaluation [16]; ISO/IEC 27000 [17]; og ISO/IEC 27002 [18] blitt benyttet. CC adresserer beskyttelse av aktiva eller verdier fra uautorisert tilgang og modifikasjon gjennom konfidensialitet, integritet og tilgjengelighet. Produktene innen datasikkerhet, spesifisert gjennom CC, kan administreres gjennom administrasjonssystemer standardisert i ISO/IEC 27000-familien. ISO/IEC 27000 og ISO/IEC 27002 presenterer et Security Management System (ISMS) som skal være et verktøy for at organisasjoner skal klare å overvåke og evaluere sine prosedyrer i forhold til skiftende omgivelser som påvirker informasjonssikkerheten. Av ETSI standarder ble det brukt en Technical Report (TR), ETSI TR 102 661, som omfatter et sikkerhetsrammeverk i rettmessig overvåkning og datalagring.

4.2 Modell over aktuell situasjon

For at det skal være enklere å se på krav til løsningen, settes det opp et scenario over hvordan situasjonen kommer til å bli dersom direktivet innføres (Figur 4.1). Scenariet inkluderer entiteter som vil påvirke løsningen på ulike måter, der alle entitetene og deres egenskaper må tas hensyn til. For at entitetene i scenariet skal samsvare med entitetene i den administrative og tekniske delen av løsningen, har vi tatt utgangspunkt i presenterte entiteter i standarden ISO/IEC 15408 [16] for å bygge opp scenariet. ISO/IEC 15408 er også kjent som Common Criteria (CC), og er standard for sertifisering av datasikkerhet. Standarden er et rammeverk der brukere av datasystemer kan spesifisere funksjonelle sikkerhetskrav, leverandører kan implementere disse sikkerhetskravene til deres produkter, og laboratorier kan deretter evaluere produktene opp mot kravene. CC tilbyr dermed en omfattende og streng standardisert prosess for å gjennomføre spesifisering, implementering og evaluering i forbindelse med produkter innen datasikkerhet.



Figur 4.1: Entiteter i løsningsscenario og forholdet mellom dem.

Entiteter i scenariet:

- Ressurser (Aktiva/Verdier): Informasjonen man ønsker å sikre (trafikkdata).
- Eier: Har rådighet over aktivaene.
- Trusselfaktor: Antatte eller faktiske faktorer med interesser for aktivaene som står i strid med eierens. Gir opphav til trusler.
- Trussel: Resultat av trusselfaktorer.
- Risiko: Grad av usikkerhet rundt eksponering av aktivaene.
- Mottiltak: Handlinger som settes i verk for å redusere risikoene.

Aktiva er de elektroniske kommunikasjonsdataene datalagringsdirektivet sier skal lagres, og informasjonen disse dataene inneholder er verdier som ønskes sikret best mulig. I tilfelle implementering av direktivet skal verken tilbydere av elektronisk kommunikasjon, politiet og påtalemyndigheter eller andre ha eierskap og direkte tilgang til dataene i følge straffeprosessloven og personopplysningsloven. *Eier* angir en enhet eller person som har godtatt å ha lederansvar for et aktivums produksjon, utvikling, vedlikehold, bruk og sikkerhet. Begrepet eier betyr dermed ikke at enheten har direkte eiendomsrett over aktivaene. Eier i dette tilfelle blir derfor både tilbydere av elektronisk kommunikasjon og samtidig et overordnet begrep for en felles interesse om vern av aktivaene. Norsk lov regulerer vilkårene for tilgang til dataene, og må brukes av eier til å definere krav for tilgang til aktivaene. Kostnadene eieren må betale for å verne om aktivaene avhenger av lagringsløsning, men faller til slutt på tilbyderne, indirekte på brukerne av kommunikasjonstjenestene, og eller delvis på myndighetene. Kostnadene for vern om aktivaene diskuteres ikke videre her. *Trusselfaktorene* kan være hackere, ondsinnede brukere, harmløse brukere som gjør feil, dataprosesser eller tilfeldigheter. Hackere eller ondsinnede brukere kan for eksempel bruke trafikkdataene samlet inn over tid til å kartlegge en persons bevegelsesmønster, kontaktnett eller utføre ID-tyveri. Trusselfaktorene utgjør reelle *trusler*.

Truslene øker *risikoene* for større usikkerhet rundt vern om aktivaene, og eieren må da iverksette *mottiltak* for å minimere disse. Mottiltak inkluderer sikkerhetstiltak både på administrativt og teknisk nivå. Gode administrative prosedyrer og god opplæring vedrørende innsamling og lagring av data, i tillegg til sterke, godt dokumenterte tekniske løsninger er viktig for best mulig beskyttelse av aktivaene.

Eieren holdes ansvarlig for aktivaene og må kunne stå for avgjørelsen om å eksponere aktivaene for truslene, og må derfor kunne forsvare de løsningene som velges for å beskytte aktivaene. For å forsvare løsningene er det viktig å kunne demonstrere at mottiltakene er tilstrekkelige og korrekte. De må være tilstrekkelige nok til at truslene kan reduseres eller stoppes og de må være korrekte ved at de fungerer slik de er ment å fungere. Den tekniske rapporten fra ETSI, ETSI TR 102 661 [19], identifiserer trussler i et overvåknings- og datalagringssystem som må tas hensyn til når en modell for datalagring settes opp:

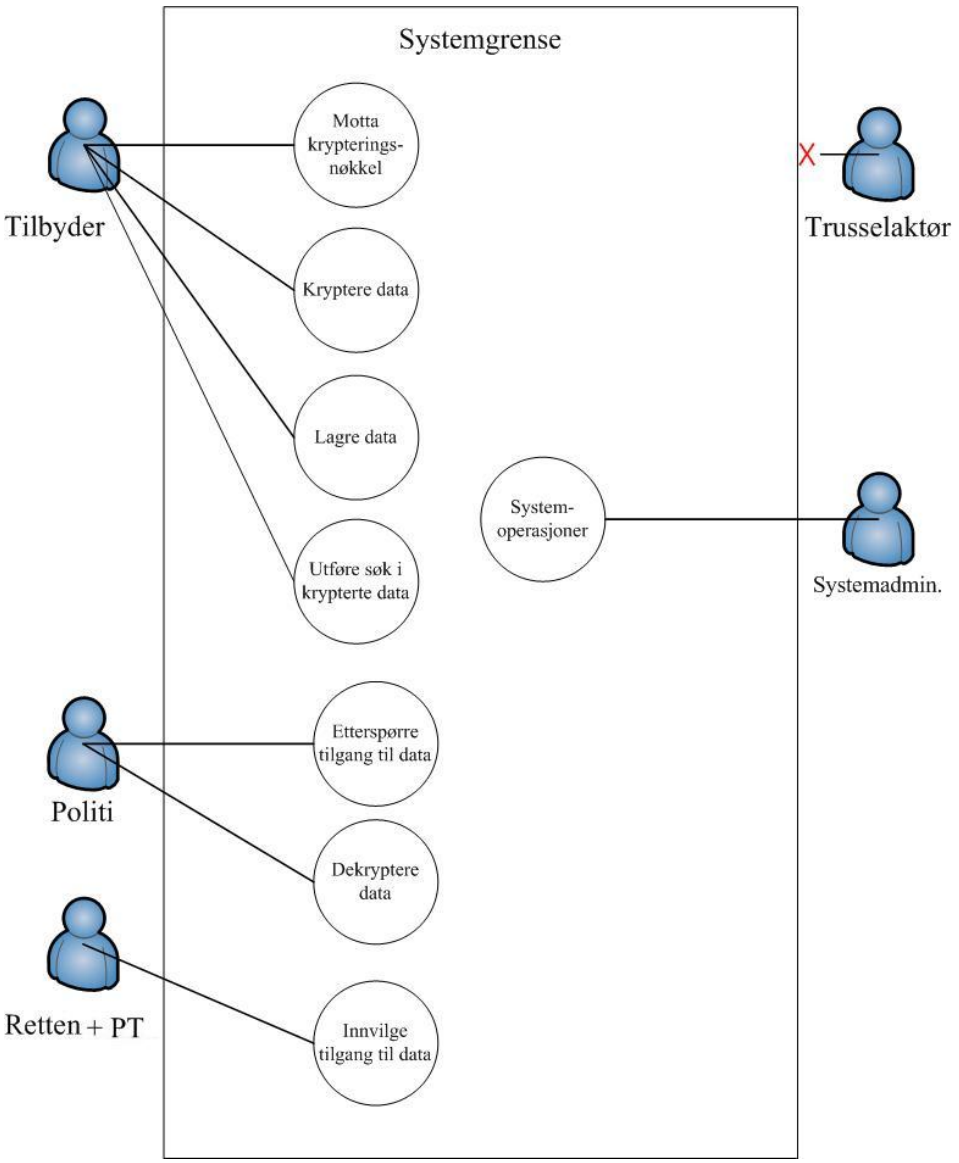
- Avsløring av ressurser (aktiva/verdier);
- Modifisering av ressurser (aktiva/verdier);
- Uautorisert aksess til ressurser (aktiva/verdier);
- Uautorisert aksess til selve datalagringssystemet eller logg-infrastrukturen;
- Misbruk av logg-infrastrukturen;
- Illegal bruk av lagret data;
- Forfalskning;
- Forlenget lagringstid;
- Gjenopprettelse av slettet data;
- Tjenestenekt;

Også mangel på tilgjengelighet, ansvarlighet og sårbarheter i nettverk, systemdesign og implementeringsprosess påpekes som mulige svakheter i et datalagringssystem.

4.3 Eksempler på bruksområder (use cases)

For å kunne se bedre hvilke krav som bør stilles til løsningen, både administrativt og teknisk, settes det opp noen brukerscenarioer. Det er tatt utgangspunkt i de identifiserte entitetene i det generelle scenariet over og det settes opp flere eksempler for bruksområder (use cases) som representerer ulike situasjoner lagringssystemet kan brukes i.

Først settes det opp et generelt use case over hvilke entiteter som kan gjøre hva i systemet (Figur 4.2). Tilbyder er eier av aktivaene (dataene) og skal kunne lagre dem, og beskytte dem med mottiltak mot trusler for å redusere risikoene. Mottiltak inkluderer først å motta krypteringsnøkkel for så å kryptere dataene. Krypteringsnøkkelen må genereres av en algoritme, så sendes til tilbyder. Tilbyder kan da kryptere dataene for å beskytte de mot trusler, og samtidig minimere risikoer forbundet med lagringen. Tilbyder skal *ikke* ha mulighet til å dekryptere eller hente ut igjen de lagrede dataene. Dette er også for å minimere risikoene for misbruk. Politiet, som er eier av aktivaene i den forstand at de skal bruke dataene, må kunne forespørre tilgang til dataene gjennom tilgang til dekrypteringsnøkkelen ved å presentere nødvendigheten av å hente ut data i forbindelse med etterforskningen. Fordi politiet må ha rettslig kjennelse og godkjenning fra PT for å hente ut lagret trafikkdata, kan de sende en forespørsel om tilgang til retten og PT. Innvilges tilgang gjennom en dekrypteringsnøkkel skal de kunne sende godkjenning av innvilgelse til tilbyder som skal søke i databasen for dem. Politiet vil så motta de riktige dataene det er innvilget tilgang til og skal kunne dekryptere disse. Dekrypteringsnøkkelen må være forskjellig fra krypteringsnøkkelen, og en algoritme for å generere asymmetriske nøkler må benyttes. Retten skal kunne motta forespørsler fra politiet, innvilge tilgang og kontrollere søket i dataene for å sikre at tilgangen er begrenset til

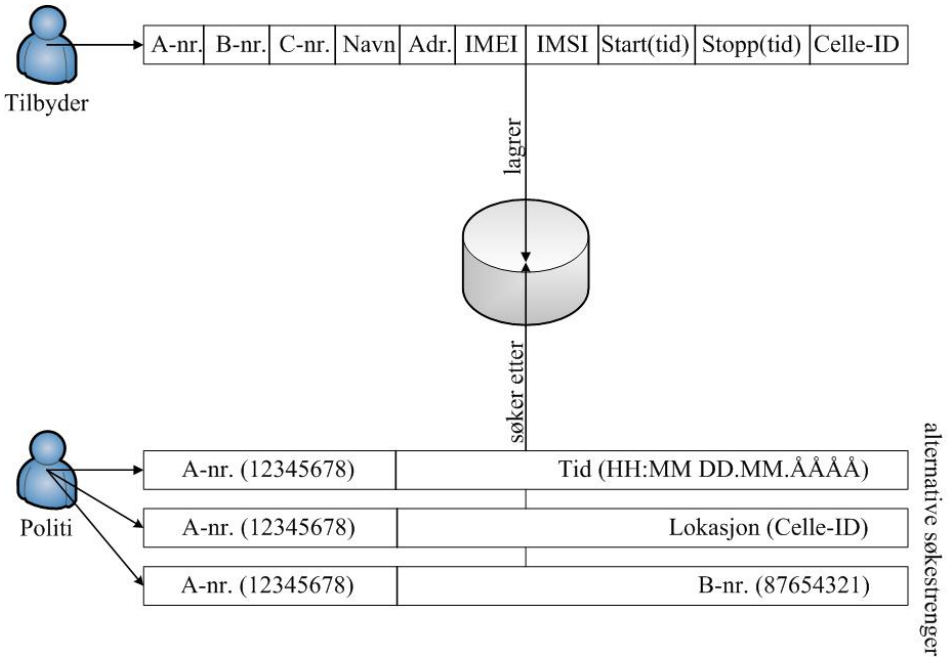


Figur 4.2: Eksempel på bruksområder i lagringssystemet.

innvilget område. Systemadministratorer skal kunne gjøre systemoperasjoner på systemet, som inkluderer vedlikehold, feilretting og gjenoppretting av hele eller deler av databasen. Trusselaktører skal på ingen måte ha tilgang til, eller kunne skaffe seg tilgang til, aktivaene.

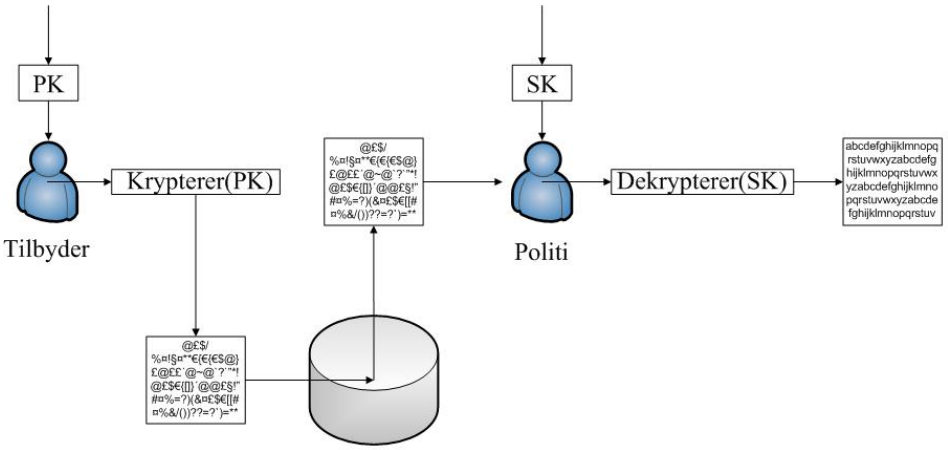
For enklere å kunne definere hvilke administrative og tekniske krav som stilles til en ny lagringsløsning for lagring av trafikkdata, ser vi på litt mer spesifikke situasjoner der lagringsløsningen er i bruk. Derfor presenteres eksempler på brukerscenarier, eller use cases.

Et eksempel på bruk er hvordan lagring og uthenting av data kan foregå (Figur 4.3). Her lagrer tilbyder alle dataene som kreves lagret i en database. Politiet har eksempelvis bare et eksakt tidspunkt og et telefonnummer; et nummer og en lokasjon; eller flere nummer, og vil hente ut disse dataene med tilhørende detaljer ved å la tilbyder søke etter dem. Kryptering og dekryptering er ikke tatt med i figuren av hensyn til det figuren skal poengtere.



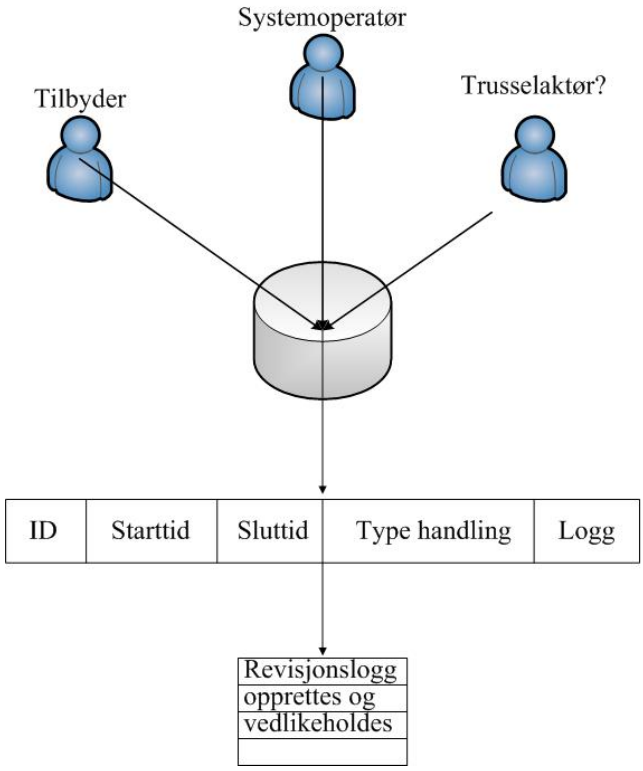
Figur 4.3: Use case for lagring av og søk på lagret trafikkdata.

Et annet brukseksempel er kryptering og dekryptering (Figur 4.4). Tilbyder skal kryptere teksten med sin mottatte nøkkel, og så lagre dataene i databasen. Politiet skal kunne motta kryptert data for så å bruke sin private nøkkel for å dekryptere teksten. Teksten skal da være i politiets besittelse i klartekst for å kunne brukes i etterforskning.



Figur 4.4: Use case for kryptering og dekryptering av data.

Det er ønskelig å logge bruken av databasens innhold, samt systemoperasjoner på databasen for og bedre kunne spore hvem som har gjort hva med dataene eller med databasen og lagringssystemet (Figur 4.5). For best mulig vern mot misbruk av aktiva (dataene) kan logging av hendelser være et bra mottiltak. Alle brukere og deres handlinger skal loggføres med ID, om man er tilbyder, politiet, retten eller systemoperatør sammen unikt brukernavn; starttid og sluttid for handlingene, gitt i hh:mm dd.mm.yyyy; type handling, om man har hentet ut data eller gjort en vedlikeholdsoperasjon; og logg, alt som leses eller skrives av data, flytting eller endringer på data eller endringer på selve systemet.



Figur 4.5: Use case for data som lagres i revisjonslogg.

Når kravene er lagt frem og løsningen presenteres, må alle brukerscenarioene slås sammen til ett system, der alle brukerscenarioene er gjennomførbare, samtidig som sikkerheten ivaretas.

Dette kapittelet har presentert hvilke standarder som har blitt brukt som grunnlag for systemkrav med informasjonssikkerhet i fokus. En modell for hvilke entiteter som inngår i løsningen har blitt satt opp, og flere brukerscenarier, eller use cases, har blitt presentert for at krav til en ny lagringsløsning lettere skulle kunne identifiseres, samt belyse hvilke funksjoner systemet bør støtte.

Kapittel 5

Krav til løsningen

Når krav til en eventuell løsning for implementering av datalagringsdirektivet skal settes opp, må noen viktige faktorer vurderes: risikoene eierne står ovenfor ved eksponering av aktivaene; juridiske, lovmessige og avtalefestede krav; og prinsipper og mål eierne (her: tilbyderne av kommunikasjonstjenestene) har utviklet for sin virksomhet. Ved risikovurderinger må tiltak for sikring av aktivaene vurderes i forhold til mulige skader ved sikkerhetssvikt. Siden direktivet fokuserer på nasjonal lovgivning ved innføringen i hvert enkelt land, må de juridiske og lovmessige bestemmelsene som gjelder i Norge for straffeprosess og personvern vurderes. Fokus på informasjonssikkerhet kan redusere risikoen for et svekket personvern samt styrke tiltroen til lagringspraksisen. Informasjonssikkerhet innebærer både god administrering rundt selve innsamlings- og lagringsprosessen, og pålitelige tekniske løsninger.

Ved formulering av de administrative og tekniske kravene for lagringsløsningen, er det tatt utgangspunkt i selve datalagringsdirektivet, norske lovtekster, standarder (seksjon 4.1) og de modellerte brukerscenariene (kapittel 4).

5.1 Administrative krav

De administrative kravene til løsningen som skal brukes ved en eventuell innføring av datalagringsdirektivet eller lignende praksis må adressere eiernes nødvendige handlinger, metoder og praksis ved mottiltak, for tilstrekkelig beskyttelse av aktivaene sett i forhold til truslene og risikoene (se figur 4.1). Slik det fremgår i høringsnotatet foreslås det ikke å opprette en sentral lagringsløsning, men at tilbyderne selv skal velge hvordan de vil administrere lagringsløsningene, enten de samarbeider seg i mellom eller har individuelle løsninger [2, Kapittel 4.7.4]. De administrative kravene ved en innføring av direktivet vil derfor i hovedsak gjelde tilbyderne av elektronisk kommunikasjon, som skal lagre dataene; men kravene må også involvere samarbeid med politiet, som eventuelt skal bruke dataene i etterforskning; og interesseorganisasjoner som er opptatte av personvern og eller informasjonssikkerhet.

Administrering inkluderer handlingene, metodene, og praksisen ved å organisere, håndtere, styre, og holde oppsyn ved aktiva. Aktivaene i dette tilfellet er trafikkdataene som skal lagres. Handlingene, metodene og praksisen settes sammen i et administrasjonssystem. Et administrasjonssystem bruker et rammeverk av ressurser slik at organisasjoner som bruker systemet skal oppnå sine mål samtidig som truslene assosiert med organisasjonens aktiva eller verdier adresseres. Systemer for administrering av informasjonssikkerhet er standardisert under International Standards Organization (ISO) og International Electrotechnical Commission (IEC), i standarden ISO/IEC 27000:2009 [17]. Systemet som presenteres her kalles Information Security Management System (ISMS). ISMS skal være et verktøy for at organisasjoner skal klare å overvåke og evaluere sine prosedyrer i forhold til skiftende omgivelser som påvirker informasjonssikkerheten. Informasjonssikkerhet har tre hoveddimensjoner i følge ISO 27000: konfidensialitet, tilgjengelighet og integritet. I forhold til en eventuell implementering av datalagringsdirektivet må konfidensialiteten være til stede gjennom kontrollert og begrenset tilgang til lagrede data; tilgjengeligheten må gjøres effektiv dersom en beslutning fra retten foreligger for uthenting av lagret data; og integriteten må bevares gjennom å hindre innbrudd og eventuell endring eller misbruk.

Administrering av informasjonssikkerhet uttrykkes gjennom formuleringen og bruken av fremgangsmåter, standarder, prosedyrer og retningslinjer innenfor informasjonssikkerhet som anvendes av alle personer (eiere) som er forbundet med aktivaene. Dersom et system for administrering av informasjonssikkerhet implementeres riktig med tanke på tilgjengelige resurser og gjeldende objektiver, vil resultatene være reduserte risikoer ved eksponering av aktivaene. En standardisert måte å implementere et slikt system på er gitt i standarden ISO/IEC 27002:2005 [18], som hører til under ISMS-familien av standarder.

5.1.1 Overordnede krav til implementeringsprosessen hos tilbyderne

Først settes det opp overordnede krav til en implementeringsløsning, så brukes disse kravene mer detaljerte krav og rutiner utarbeides innenfor personellsikkerhet, fysisk sikkerhet, prosedyrer, aksesskontroll, kryptering og samsvar med gjeldende lovverk. Disse presenteres derfor påfølgende etter de overordnede kravene. Risikovurderinger og kontinuitetsplanlegging er også viktig for en ryddig implementeringsprosess av systemer for informasjonssikkerhet i følge ISO/IEC 27002 [18].

For å sette opp administrative krav til dette arbeidets forslag til lagringsløsning, analyseres viktige suksessfaktorer for en vellykket implementering av et administrasjonssystem for informasjonssikkerhet i standarden ISO/IEC 27000 og understandarden ISO/IEC 27002, og anbefalinger fra ETSI TR 102 661 vedrørende overvåknings- og datalagringssystemer:

- informasjonssikkerhetsprinsipp, -mål og -aktiviteter bør gjenspeile tilbydernes målsetninger;
- rammeverk for iverksettelse, vedlikehold, overvåkning og forbedring av informasjonssikkerhet bør være i samsvar med tilbydernes kultur;
- synlig støtte og forpliktelse fra toppledelsen hos tilbyder og hos politiet;

- god forståelse for informasjonssikkerhetskrav, risikovurdering og risikostyring;
- veiledning for effektiv markedsføring av informasjonssikkerhet ovenfor ledere og ansatte hos tilbyder og hos politiet, og hos andre interessenter for god bevisstgjøring;
- informasjon til ledere og ansatte hos tilbyder og hos politiet, og andre interessenter om informasjonssikkerhetsstandarder og -praksis;
- tilrettelegging av tiltak for håndtering av informasjonssikkerhet;
- hensiktsmessig bevisstgjøring, utdanning og opplæring;
- etablering av en effektiv prosess for håndtering av informasjonssikkerhetsbrudd;
- iverksettelse av et målesystem for evaluering av administrasjon av informasjonssikkerhet samt tilbakemeldinger med forslag til forbedringer;

Personellsikkerhet

De viktigste kravene ved personellsikkerhet gjelder å sikre at administrasjonen og hver enkelt ansatt hos både tilbydere, politiet, PT og retten er forstår sitt ansvar og reduserer risikoen for tyveri, svindel eller misbruk av utstyr. Ivaretakelse av sikring av trafikkdata bør prioriteres ved ansettelse, og ved endring av stillingsbeskrivelse eller ansvarsområder eller avslutning av ansettelse. Det bør foreligge en formell prosess for endring av rettigheter i forhold til posisjon hos tilbyder, politiet eller retten. Prosessen bør inkludere bevisstgjøring og regelmessig ajourføring av tilpassede retningslinjer innenfor informasjonssikkerhet som er relevante for hver enkelts nåværende jobbfunksjon. Det kan være gunstig å opprette et team, eller en gruppe mennesker, som har som ansvarsområde og utføre alle handlinger som har med datalagringssystemet og gjøre [19, Kapittel 7]. Krav for god personellsikkerhet bør derfor inkludere:

- at ansatte med tilgang til lagringsløsningens mekanismer begrenses til et minimum;

- at alle ansatte forstår sitt ansvar for å redusere risiko forbundet med aktiva:
  - sikkerhetsansvar bør spesifiseres i arbeidskontrakt og stillingsbeskrivelser;
  - søkere til jobber der sikkerhet spiller en stor rolle bør vurderes nøye;
  - alle ansatte bør undertegne en kontrakt som inkluderer taushetsplikt;
  - ledelsen må sørge for at sikkerheten ivaretas under hele perioden en person er ansatt;
- at alle ansatte er klar over trusler og risiko forbundet med aktiva;
- spesifiering av prosesser for å håndtere brudd på sikkerhetsreglene;
- prosesser for sletting eller endring av en ansatts rettigheter i forhold til behandling av eller tilgang til aktiva;

Bedrifter som jobber med sensitiv og hemmelig informasjon har ofte lignende rutiner vedrørende personellsikkerhet fra før. Gjeldende rutiner kan muligens videreføres ved innføring av datalagringsdirektivet eller lignende ny lagringspraksis. Rutiner for personellsikkerhet nevnes ikke ytterligere her på grunn av dette arbeidets begrensede omfang og fokus, men interesserte lesere henvises til en grundigere utredning av krav til personellsikkerhet i standarden for administrasjon av informasjonssikkerhet ISO/IEC 27002 [18, Kapittel 8].

Fysisk sikkerhet

For å forhindre uautorisert tilgang til eller skade på aktiva/elektroniske trafikkdata, er den fysiske sikkerheten viktig. Utstyret som holder på trafikkdataene bør plasseres i sikre områder beskyttet av sikkerhetssoner med tilpassede sikkerhetsbarrierer og adgangskontroll. Miljømessige trusler bør også tas hensyn til. Det bør beskyttes mot naturlige trusler som brann,

oversvømmelse, eksplosjoner, strømbrudd eller fremkalte ulykker. Krav til fysisk sikkerhet rundt datalagring bør inkludere:

- etablering av fysiske sikkerhetssoner for beskyttelse av områder der dataene blir lagret;
- sikre områder bør beskyttes med adgangskontroll med adgang kun for autorisert personell;
- fysisk sikkerhet for kontorer, rom og utstyr bør planlegges og innføres;
- planer for beskyttelse mot og håndteringen av miljømessige trusler bør utarbeides og innføres;
- retningslinjer for arbeid utført i sikre områder bør utformes og innføres;
- utstyr bør beskyttes mot strømbrudd;
- kabler for strøm og kommunikasjon bør beskyttes mot avlytting og skade;
- utstyr bør vedlikeholdes for tilgjengelighet og integritet;

Det er utenfor dette arbeidets omfang å spesifisere kravene fysisk sikring og ytterligere. Ytterligere krav finnes i ISO/IEC 27002 [18, Kapittel 8].

Prosedyrer

Å sikre korrekt og sikker drift av informasjonsbehandlingsutstyr er viktig for god administrasjon av informasjonssikkerhet. Ansvar og prosedyrer for administrasjon og drift av alt utstyr som brukes i den presenterte løsningen for ny, utvidet datalagring bør etableres hos både eier (her, tilbyder) og politiet, PT og retten. For å redusere faren for utilsiktet eller overlagt misbruk av lagringsløsningen bør hensiktsmessige driftsprosedyrer utvikles og innføres. Driftsporsedyrene bør inkludere prosedyrer for lagring av data, rutiner for opprettholdelse av integritet og tilgjengelighet av de lagrede dataene gjennom sikkerhetskopiering, regler for informasjonsutveksling samt overvåkning

av informasjonsbehandlingsaktiviteter gjennom revisjonslogging [18].

Prosedyrer for lagringen kan inkludere:

- håndtering og merking av alle data etter satt sikkerhetsklassifisering;
- aksessrestriksjoner slik at uautorisert personale ikke får tilgang;
- vedlikehold av register over autoriserte brukere;
- å sikre at inndata er komplett;
  - kontroll av verdiene dataene inneholder, at de er innenfor grenseverdiene både i verdi og volum;
  - ingen ugyldige tegn;
  - ingen manglende eller ufullstendige data;
  - periodisk gjennomgåelse av innhold i nøkkelfelter;
  - alle endringer i inndata må være autorisert;
  - prosedyrer for å teste om inndata er sannsynlige;
  - definert ansvar for alle medarbeidere som er involvert i lagringsprosessen;
  - opprettelse av logg for handlinger i forbindelse med inndataene (se avsnitt om revisjonslogger);
- å sikre at utdata er komplett og godkjent som utdata;
- lagring i et miljø som er tilpasset sikkerhetsklassifiseringen;
- kryptering av data med en sterk kryptografisk algoritme tilpasset sikkerhetsklassifiseringen;
  - administrasjon av kryptografiske nøkler bør opprettes for håndtering av gjenopprettelse av informasjon i tilfelle tap av kryptografiske nøkler;
- begrense distribusjon av data til et minimum;

- tydelig merking av alle kopier av dataene med navn på bruker som skal ha kopien;
- regelmessig gjennomgang av distribusjonslister og register over autoriserte brukere med tilgang;
- sletting av informasjon som skal fjernes på grunn av utgått lagringstid, og sikring av at disse ikke skal kunne gjenopprettes;
- rapportering av brudd på informasjonssikkerheten;

Disse prosedyrene gjelder informasjonen i de dataene som skal lagres i henhold til datalagringsdirektivet, både i lagringsmedium og under all annen form for distribusjon av informasjonen.

For å opprettholde integriteten og tilgjengeligheten av informasjonen i de lagrede dataene bør det etableres rutiner for sikkerhetskopiering [18]. Sikkerhetskopiering bør tas og testes regelmessig for å sikre at informasjonen kan gjenopprettes etter datasammenbrudd eller feil. Omfanget og frekvensen av sikkerhetskopieringen i dette tilfellet bør gjenspeile hvor kritisk informasjonen er for vellykket etterforskning og forhindring av alvorlig kriminalitet, hvilket muligens bør spesifiseres ytterligere. Rutiner rundt sikkerhetskopieringen kan inkludere:

- definisjon av nødvendig hyppighet av sikkerhetskopiering;
- opprettelse av fullstendige registre for sikkerhetskopierte data;
- opprettelse av dokumenterte prosedyrer for gjenoppretting;
- lagring på lokasjoner med tilstrekkelig fysisk avstand fra lokasjon for hovedlagringen slik at ulykker på hovedlokasjon ikke påvirker sikkerhetskopiene;
- anskaffelse av de samme fysiske beskyttelsene for sikkerhetskopiert data som for originale data;

- anskaffelse av de samme tekniske beskyttelsene for sikkerhetskopiert data som for originale data, inkludert kryptering av data;
- test av sikkerhetskopiert data for å sikre deres tilgjengelighet og integritet;
- test av gjenopprettelsesprosedyrer for å sikre at de er effektive nok;

Regler for informasjonsutveksling skal beskytte informasjonen som utveksles med eksterne enheter, og fastsette utvekslingsavtaler som skal være i samsvar med norsk lovgivning. Prosedyrene rundt informasjonsutveksling bør beskytte selve informasjonen og medier eller kommunikasjonsutstyr som inneholder informasjonen under forsendelse. Prosedyrene som opprettes for informasjonsutveksling bør ta hensyn til gjeldende prosedyrer for beskyttelse av informasjonen av hensyn til personvern inkludert krypteringsteknikker, samt gjeldende norsk lovgivning for utlevering av trafikkdata [5], [7]. Utleveringsavtaler mellom tilbydere, PT, retten og politiet bør opprettes og kan inkludere:

- prosedyrer for å varsle avsender om sending, overføring og mottak;
- prosedyrer for å sikre sporbarhet;
- tekniske minimumsstandarder for pakking og sending;
- avtaler om deponering av programvare;
- standarder for identifisering av kurér;
- ansvar og forpliktelser i forbindelse med tap av informasjon;
- bruk av avtalt merking av sikkerhetsklassifisering for tilstrekkelig beskyttelse;
- eierskap til og ansvar for beskyttelse av data;
- tekniske standarder for skriving og lesing av informasjon og programvare;

- nødvendige spesialtiltak for beskyttelse av sensitivt materiale, inkludert krypteringsnøkkel;

Avtalens sikkerhetsinnhold bør gjenspeile graden av sensitivitet og viktighet informasjonen som skal utveksles har [18].

Overvåkning av informasjonsbehandlingsaktiviteter kan gjøres ved bruk av revisjonslogger. Alle hendelser som gjøres i forbindelse med lagringen av trafikkdata, inkludert brukeraktivitet, unntak og informasjonssikkerhetshendelser, bør logges for å avsløre uautoriserte informasjonsbehandlingsaktiviteter. Operatørlogger og feillogging, eller revisjonslogger, bør benyttes for å identifisere problemer i lagringssystemet, og aksesskontrollen bør overvåkes. Revisjonsloggene bør i følge standarden for administrasjon av informasjonssikkerhet [18] inneholde:

- brukernavn (unikt for hver bruker);
- dato, tidspunkt og type hendelse;
- lokalisering hvis mulig;
- oversikt over vellykkede og avviste forsøk på systemtilgang;
- oversikt over vellykkede og avviste forsøk på tilgang til lagrede trafikkdata;
- endringer i systemkonfigurasjon;
- bruk av privilegier;
- filer som aksesseres og typen tilgang;
- nettverksadresser og protokoller;
- eventuelle alarmer som utløses av aksesskontrollsystemet;
- aktivering eller deaktivering av beskyttelsessystemer, inkludert antivirussystemer eller deteksjonsprogrammer for innbrudd;

Siden disse loggene inneholder brukernavn, lokasjon, tid og dato bør disse loggene også beskyttes av hensyn til personvernet, og holdes konfidensielle. Loggene bør beskyttes mot manipulering og uautorisert tilgang, og selv ikke systemadministratorer bør ha mulighet til å endre eller slette loggene.

Aksesskontroll bør reguleres gjennom administrasjon av brukertilgang. Det mest elementære alle med permanent eller midlertidig tilgang til lagrede data kan gjøre er å avslutte aktive sesjoner når de er ferdige, logge av maskinger brukt til aksess av lagringssystemet, samt å sikre at pc-er eller terminaler som brukes hindrer uautorisert bruk med tastaturlås eller passordtilgang når de ikke er i bruk. Dersom noen skulle skrive ut informasjonen dataene inneholder, bør de alltid være lagret på et tilstrekkelig sikkert sted, og papirer og utskiftbare lagringsmedier bør alltid være ryddet fra personlige arbeidsplasser. Sesjoner der tilgang til dataene er gitt bør også begrenses av oppkoblingstid og med tidsavbrudd ved inaktivitet. Brukertilgang bør kontrolleres på grunnlag av sikkerhetsbehovene rundt de lagrede trafikkdataene, samt at relevant lovgivning om beskyttelse av og tilgang til trafikkdataene må tas hensyn til. Formelle prosedyrer bør foreligge for å kontrollere tildeling av aksessrettigheter til dataene og systemene som kontrollerer lagring og uthenting av disse. Aksesskontrollprosedyren bør forutsette:

- unike brukernavn, slik at hver enkelt bruker med tilgang kan stilles ansvarlig for sine handlinger;
- kontroll av at bruker har autorisasjon fra det norske rettssystemet til å hente ut informasjon fra de lagrede dataene;
- kontroll av tildelt aksessnivå, slik at brukere kun gis tilgang til de tjenester og data de uttrykkelig er autorisert til;
- begrense antallet tillate misslykkede aksessforsøk til systemet;

Det vil være nødvendig med kontinuerlig gjennomgang av ulike brukeres aksessrettigheter for opprettholdelse av en effektiv aksesskontroll. I vårt tilfelle er politiets tilgang til disse historiske elektroniske kommunikasjonsdataene

regulert av Ekomloven [5] og Straffeprosessloven [7] (se seksjon 2.3.2). Politiet må altså ha rettslig kjennelse for å få utlevert data, og i tillegg skal Post- og Teletilsynet godkjenne at taushetsplikten til eier av dataene må overskrides. De enkeltpersonene hos både politiet og hos tilbyder som henter ut data bør likevel identifiseres ved unike brukernavn og personer involvert fra rettsvesenet og Post- og Teletilsynet bør også kunne identifiseres. Hensiktsmessige autentiseringsmetoder bør altså implementeres for å kontrollere tilgang for eksterne forbindelser. Eksterne forbindelser bør heller ikke være lov fra hvilken som helst lokasjon eller fra hvilket som helst utstyr. Slike autentiseringsmetoder kan utdypes gjennom valg av teknologiske løsninger (se seksjon 5.2 og 7.2).

I og med at tilbydere, politiet, retten, og Post- og Teletilsynet allerede har erfaring med lagring av data regulert av personopplysningsloven, straffeprosessloven og Ekomloven, eksisterer det allerede prosedyrer vedrørende lagringspraksis og tilgang til lagrede data. Mange av de nevnte administrative kravene over praktiseres kanskje allerede, og kan derfor bygges videre på. Det er ikke et krav at alle disse prosessene settes i gang på nytt, men da overføres til den nye, separate lagringspraksisen og -løsningen.

Kravene presentert i denne seksjonen gjelder tilbyders administrative praksis ved innføring av datalagring med ny, separat lagringsløsning. Tilbyder av elektroniske kommunikasjonstjenester må administrere innføringen av en ny lagringsløsning samtidig som de må administrere et godt samarbeid med politiet som eventuelt skal bruke dataene, og interesseorganisasjoner som vil beskytte de lagrede dataene fra de trusler som oppstår ved den nye lagringspraksisen.

5.1.2 Krav til samarbeid ved implementeringsprosessen

I sammenheng med datalagringsdirektivet bør det sørges for hensiktsmessig kontakt med interessegrupper for å identifisere krav til beskyttelse av sensitiv informasjon. Interessegrupper i vårt tilfelle er personvernorganisasjoner som vil verne om de lagrede dataenes tilgjengelighet, integritet og ko-

rrekthet, eller sikkerhetsspesialister som vil vedlikeholde sikkerheten rundt lagringen. Kontakten med aktuelle interessegrupper bør oppnå:

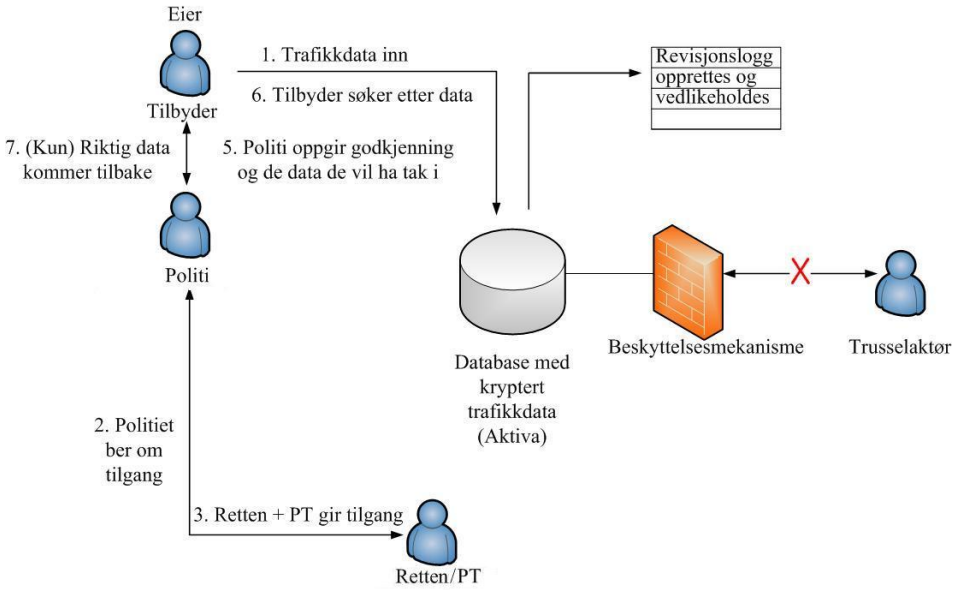
- forbedret kunnskap om informasjonssikkerhet og personvern;
- økt forståelse av informasjonssikkerhet og personvern;
- tidlig varsling, veiledning og korrigering i forbindelse med svekkelse av personvernet og sårbarheter ved informasjonssikkerhet;
- tilgang til råd om informasjonssikkerhet og personvern fra spesialister;
- kontinuerlig utveksling av informasjon om trusler;
- at egnede kontaktpersoner kan håndtere brudd på personopplysningsloven eller informasjonssikkerheten under lagringen;

I denne seksjonen har det blitt lagt frem krav til administrative prosedyrer rundt en ny datalagringsløsning. Mange av disse prosedyrene og tiltakene eksisterer allerede hos noen tilbydere, men det er ikke sikkert alle tilbyderne er like seriøse rundt alle tiltakene. Administrative overordnede krav for implementeringen, personellsikkerhet, fysisk sikkerhet, og overordnede prosedyrer for lagring, aksesskontroll, sikkerhetskopiering utleveringsavtaler og opprettelse av revisjonslogger har blitt presentert.

5.2 Tekniske krav

Datalagringsdirektivet spesifiserer at de lagrede dataene skal brukes til etterforskning, avsløring og rettsforfølgelse av alvorlig kriminalitet, og skal kun utleveres til kompetente nasjonale myndigheter i overensstemmelse med nasjonal lovgivning (se seksjon 2.1). Dette betyr at tilbydere som lagrer dataene må ha kontakt med politiet angående tilgang til dataene. Som nevnt (i seksjon 2.3.2) kan politiet etterspørre informasjon om abonnenten, for eksempel hvilken abonnent som tilhører et telefonnummer, uten rettslig kjennelse, men de må ha tillatelse fra retten hvis de skal ha historiske trafikkdata utlevert.

Situasjonen som spesifiserer kravene til den tekniske løsningen kan sees på som et samarbeid mellom entitetene tilbyder, politiet, PT, retten og trusselaktør, og deres forhold til dataene som skal lagres. Dataflyten mellom entietene regulert av norsk lovgivning bestemmer de tekniske kravene til løsningen (Figur 5.1).



Figur 5.1: Dataflyt mellom entiteter med tilknytning til lagret trafikkdata.

Krav som stilles til den tekniske løsningen inkluderer:

- Tilbyder skal kunne legge inn data, men ikke hente ut;
  - Inndata må være korrekt (jf. seksjon 5.1.1);
- Lagrede data bør krypteres under lagring;

- Politiet må ha en form for godkjenning fra retten og PT for å hente ut data;
- Politiet skal ikke gis tilgang til data i hele databasen, kun de data de har fått godkjenning til å bruke i etterforskning;
  - Utdata må være korrekt (jf. seksjon ref subsubsec: prosedyrer);
- Søk i databasen etter forespurt data bør være så effektivt som mulig;
  - Nøyaktige tidspunkter må kunne gjenkjennes i registrerte tidsrom (range-queries);
  - Det må være mulig å søke på flere kriterier samtidig;
- Beskyttelsesmekanismer for uautorisert tilgang til dataene bør iverksettes;
- Revisjonslogger bør opprettes kontinuerlig ved utførte handlinger som berører dataene;
- Et begrenset utvalg av utstyr skal kunne oppnå tilgang;

Tilbyder skal ikke automatisk ha tilgang til å hente ut lagret data, da de skal ha egne lagrede data som kan brukes til fakturering. Tilbyderne skal kun kunne lagre data, og dataene bør verifiseres før de krypteres og lagres. Tilbyder skal kun søke i databasen når en rettslig beslutning er gitt, og politiet skal ikke kunne spørre etter lagrede data uten godkjenning av norsk rettsvesen og PT. Når tillatelse er gitt politiet til å bruke de dataene de har forespurt tillatelse til å bruke, skal uthentingen fra tilbyder være effektiv og nøyaktig. Kun de dataloggene det er gitt tillatelse til å hente ut skal hentes ut. Verken politiet eller tilbydere skal ikke kunne søke fritt i databasen og lete etter dataene, og en sjekk som bekrefter korrektheten av dataene er nødvendig. Det må forøvrig være mulig å søke i tidsspekter, og ikke bare på nøyaktige tidsrom. I loggene vil IP-adresser kanskje være tilknyttet brukere ved på- og avloggingstidspunkt, og et nøyaktig tidspunkt politiet vil ha informasjon om vil da ligge et sted i dette tidsrommet. Det

er derfor et krav at man skal kunne finne i hvilket tidsspekter dette nøyaktige tidspunktet befinner seg, for eksempel om $t \in [t_1, t_2]$. Det skal også være mulig og samtidig søke på telefonnummer $n \in [n_1, n_2]$, og IP-adresser $a \in [a_1, a_2]$. Et søk etter data skal kunne tilfredsstille alle søkekriteriene samtidig. Øvrige beskyttelsesmekanismer, som et begrenset antall brukere og utstyr på tilgangslisten samt revisjonslogger, bør iverksettes for å forhindre uautorisert tilgang. Det bør ikke kunne benyttes PC-er på offentlige steder (type internettcafé) eller private hjemmenettverk til å hente ut data eller gjøre systemoperasjoner, og revisjonslogger bør opprettes kontinuerlig for å overvåke bruken av de lagrede dataene.

I forbindelse med å beskytte informasjonens konfidensialitet, autentisitet og integritet, kan kryptografi benyttes. Ved å kryptere informasjonen som lagres kan man oppnå konfidensialitet, integritet og ikke-benektelse. Sensitiv informasjon kan beskyttes ved å kryptere den, digitale autentiseringskoder kan benyttes for å bevare integriteten, samt at man ved hjelp av kryptografiske teknikker kan bevise at en hendelse har funnet sted eller ikke og dermed oppnå ikke-benektelse. Ved bruk av kryptografi, må det tas hensyn til:

- påkrevd beskyttelsesnivå basert på risikovurdering;
- styrke og kvalitet på valgt krypteringsalgoritme;
- virkningen av kryptografien som velges, og at den beskytter ønskede områder på ønsket nivå;
- at valgt kryptografi håndterer eksakte søk og søk på et spekter (mtp tidsrom);
- at det skal kunne søkes på informasjonen som er lagret og kryptert;

For å støtte bruken av kryptografiske teknikker brukes det kryptografiske nøkler. Utstyret som benyttes til å generere eller lagre nøklene bør beskyttes fysisk. For økt sikkerhet bør nøklene ha en begrenset tidsperiode der de er gyldige som nøkler. Forespørsler fra retten, PT og politiet om tilgang

til kryptografiske nøkler må kunne håndteres for å kunne bruke kryptert informasjon i ukryptert form under etterforskning eller i rettssaker.

Kapittel 6

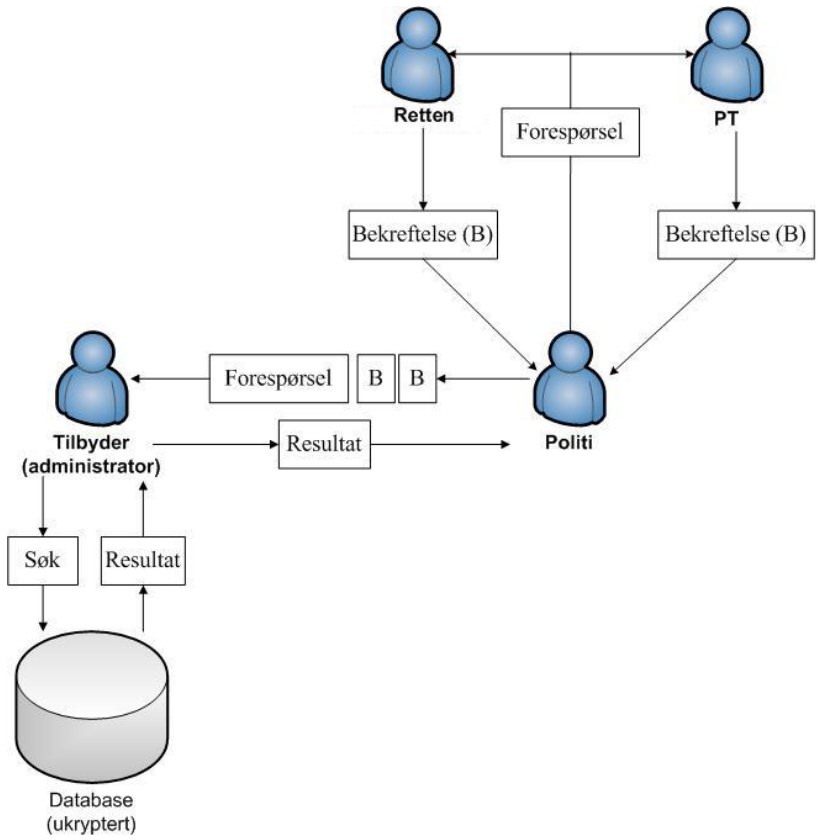
Mulige kryptoalgoritmer

Det har vært foreslått flere typer løsninger og videreutvikling av tidligere løsninger når det gjelder søk i, tilgang til og distribusjon av krypterte data [20], [21], [22], [23], [24]. I denne seksjonen presenteres noen av de kryptoalgoritmene som kan brukes til den tekniske løsningen som kan besvare problemstillingen til dette arbeidet på en best mulig måte.

6.1 Ukryptert database

Det enkleste alternativet når det kommer til kryptoalgoritmer, er og ikke bruke en kryptoalgoritme i det hele tatt. Bruk av ukryptert database er tilsvarende med dagens praksis for lagring av elektroniske kommunikasjonsdata. Dette betyr at tilbyder kan se bort fra å kryptere dataene som lagres, og bare lagre dataene i klartekst. Aksesskontrollen blir da hovedmomentet for informasjonssikkerheten. Et begrenset antall brukerkontoer med administratorrettigheter og få brukere med full tilgang til de ukrypterte dataene, kan i utgangspunktet være en tilstrekkelig lagringsløsning for noen tilfeller. Da trenges det ingen spesiell eller tilpasset kryptoalgoritme for å spørre eller søke etter data som skal utleveres i forbindelse med etterforskning. Politiet sender altså bare over den informasjonen de har og hva de vil søke etter sammen med kjennelsesdokumentet fra retten, til tilbyder. En bruker hos

tilbyder med riktige rettigheter søker så i databasen etter riktig person, IP-adresse eller telefonnummer. De dataene i søket som passer til politiets forespørsel sendes deretter over til politiet (Figur 6.1). Eventuelt kan dataene sendes kryptert eller over sikre kanaler.



Figur 6.1: Mulig lagringsløsning med ukryptert database.

6.2 Identitetsbasert kryptering

Identitetsbaserte kryptosystemer (IBE) [20] ble presentert allerede i 1985. Teorien baserer seg på at det opereres med et lukket system, som for eksempel mellom styremedlemmer i et multinasjonalt selskap eller grener av en stor bank. Øverste organ i systemet må være til å stole på, og dette organet genererer nøklene som brukes i kryptosystemet. Hver bruker tildeles et personlig smartkort når brukeren kobler seg til nettverket eller systemet første gang. Informasjonen i kortet lar brukeren signere og kryptere dataene som lagres eller sendes, og dekryptere og verifisere meldinger som mottas. Når en bruker A vil gjøre data tilgjengelig *kun* til bruker B, krypterer A meldingen med Bs navn og nettverksadresse. Ved mottakelse av meldingen dekrypterer B meldingen med den skjulte nøkkelen i hans smart kort. Nøklene må genereres av en uavhengig algoritme, slik at brukerne selv ikke kan generere dem. Genereringsalgoritmen må være privilegert, og ha kunnskap om noe ingen av brukerne har (for eksempel faktorisering av et stort tall som i RSA), slik at den kan generere nøkler til alle brukerne. Bruker B kan dekryptere alle data som er kryptert med sitt navn og sin nettverksadresse.

6.3 PECKS og HVE

Public Key Encryption with Conjunctive Keyword Search (PECKS) [21] inkluderer bruk av asymmetriske nøkler, en public key, en felles/offentlig nøkkel, og en secret key, skjult/personlig nøkkel. Dette passer det aktuelle scenariet, slik at tilbyder krypterer med politiets felles nøkkel, men dekrypteringen kan bare skje med politiets personlige nøkkel. Slik får ingen andre enn politiet mulighet til å lese dataene i klartekst. PECKS er en kryptoalgoritme som bruker predikater [25]. Kryptoalgoritmer som bruker predikater assosierer hver krypterte tekst, C_t, med en binær vektor med attributter $\mathbf{x} = (x_1, ..., x_n)$, og assosierer nøkkelen K med predikatet P [22]. Nøkkelen K kan kun dekryptere en kryptert tekst C_t hvis vektoren med attributter tilfredsstiller predikatet til nøkkelen. Rammeverket fra [25] brukes for forespørsler (queries) på krypterte data:

Vi lar $\sum$ være et endelig sett av binære strenger. Et predikat P over $\sum$ er en funksjon slik at $P : \sum \rightarrow \{0,1\}$. Vi sier at $I \in \sum$ tilfredsstiller predikatet dersom $P(I) = 1$. Predikater kan sees på som delvise dekrypteringsnøkler som avslører delvise mengder klartekst, nok til å finne ut om man har funnet de dataene man søker etter. Φ defineres til å være et sett med predikater over $\sum$. Et Φ-søkbart asymmetrisk nøkkelsystem består da av følgende algoritmer [22]:

- ***Setup***(λ): en algoritme som tar en sikkerhetsparameter inn og gir en felles nøkkel PK og en skjult/personlig nøkkel SK ut.
- ***Kryptér***(PK, I, M): krypterer klartekst (I, M) ved bruk av felles nøkkel PK. $I \in \sum$ er det søkbare feltet, og $M \in \mathcal{M}$ er data.
- ***Generer symbol***(SK, $\langle P \rangle$): tar den skjulte nøkkelen SK og beskrivelsen av predikatet $P \in \Phi$ som input. Gir ut et symbol (et "token") TK_P
- ***Forespørsel***(TK, C_t): tar symbolet TK_P for et predikat $P \in \Phi$ og kryptert tekst C_t som input, og gir ut meldingen $M \in \mathcal{M}$ eller $\perp$ (ingenting).

Dette betyr at dersom C_t er en kryptering av (I, M) gir algoritmen ut M når $P(I) = 1$, og $\perp$ ellers. I den aktuelle situasjonen er det ikke ønskelig at meldingen M skal komme ut i klartekst. De ulike forespørslene lages ut fra symbolene (tokens) som igjen er generert ved hjelp av den skjulte nøkkelen, som i det aktuelle tilfelle kun besittes av politiet. Trusselaktøren har dermed ingen mulighet til å generere forespørsler og kjøre angrep med disse. Det er heller ønskelig at algoritmen for *Forespørsel* kun skal si *om* C_t er en kryptering av (I, M), og gi ut sann (true) isteden, slik at man vet at forespørselen stemmer. Dersom forespørselen ikke stemmer, er det ønskelig å få ut $\perp$, og det vil uansett aldri være ønskelig å få tilgang til meldingen i klartekst som et svar på søk. Det er viktig å bemerke seg at i dette systemet er det ingen dekrypteringsalgoritme som kan dekryptere M. Selv om ikke

dekryptering er nødvendig for å kunne søke i kryptert data, er det ønskelig i den aktuelle situasjonen at man skal kunne dekryptere meldingen M på samme måte etter man har mottatt at man har med riktig C_t og gjøre ($P(I) = 1$). Vi definerer derfor:

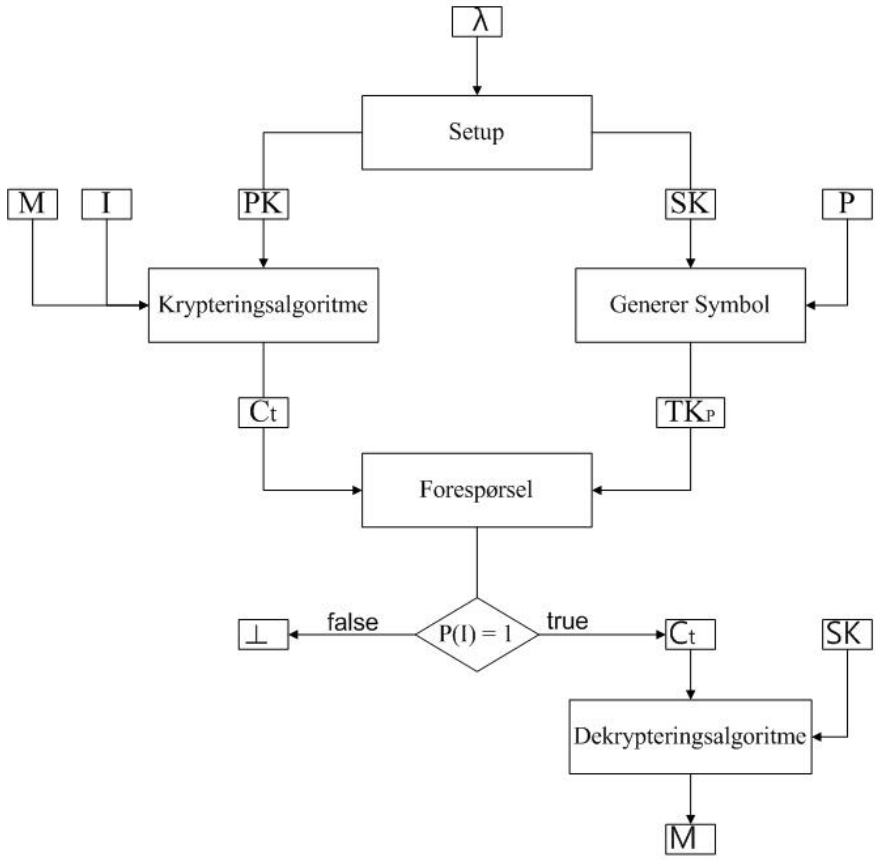
- ***Dekryptér***(SK, C_t): tar inn den skjulte nøkkelen og den krypterte meldingen C_t. Gir ut (M).

Denne algoritmen er ikke en nødvendig del av selve søkeprosessen, men den er nødvendig for å kunne lese de krypterte dataene fra søkeresultatet. Alle algoritmene henger nøye sammen, og de er avhengige av hverandres input for å fungere som et kryptosystem som eventuelt kan brukes i det aktuelle tilfellet til kryptering og dekryptering, søk i krypterte data, og aksesskontroll (Figur 6.2).

Når vi nå har definert et Φ-søkbart krypteringssystem, kan Hidden-Vector Encryption (HVE) [22] brukes både for aksesskontroll og for å oppnå et mer effektivt søkesystem tilpasset flere ulike måte å søke/forespørre på. Ulike søk inkluderer søk på flere samsvarende forespørsler samtidig, forespørsler om delmengder, eller innenfor et gitt spekter av verdier. HVE er en underalgoritme av PECKS, og er derfor også en type kryptoalgoritme som bruker predikater [22]. I HVE er predikatet en binær vektor $\mathbf{y} = (y_1, ..., y_n)$ som også inkluderer likegyldige verdier, representert ved *. Nøkkelen K kan bare dekryptere C_t dersom $\mathbf{x}$ og $\mathbf{y}$ stemmer for alle i der $y_i \neq *$. Predikatet i HVE opprettes ved å la $\sum$ være en endelig mengde, og * et symbol som ikke er i $\sum$. Vi definerer $\sum_* = \sum \cup \{*\}$. Symbolet * representerer altså her en likegyldig verdi. I forespørsler etter en delmengde eller innenfor et spekter settes $\sum = \{0, 1\}$, som i PECKS. For $\sigma = (\sigma_1, ..., \sigma_l) \in \sum_*^l$ defineres predikatet:

$$P_\sigma^{HVE}(x) = \begin{cases} 1 & \text{h}vis\ for\ alle\ \ i = 1, ..., l : (\sigma_i = x_i \quad eller \quad \sigma_i = *) \\ 0 & ellers \end{cases}$$

Dette betyr at vektoren $\mathbf{x}$ er lik σ på alle punkter der $\sigma \neq *$.
Da er $\Phi_{HVE} = P_\sigma^{HVE}\ for\ alle\ \sigma \in \sum_*^l$, der l er bredden av HVE.



Figur 6.2: Hendelsesforløp og sammenheng mellom PECKS-algoritmene.

HVE spesifiserer hvordan algoritmene *Setup*, *Kryptér* og *Dekryptér* beregner krypteringsnøkler, og hvordan disse brukes til å kryptere eller dekryptere [25]. Vi lar først $\sum = \mathbb{Z}_m$ for et heltall m. Vi setter $\sum_* = \mathbb{Z}_m \cup \{*\}$. Vi setter M til å være en undermengde av $\mathcal{M}$ av $\mathbb{G}_T$, altså $|\mathcal{M}| < |\mathbb{G}_T|^{1/4}$. Så spesifiseres algoritmene:

Setup(λ): velger først to tilfeldige primtall p, $q > m$ og lager en bilineær gruppe $\mathbb{G}$ i sammensatt orden av $n = p, q$. Så velges tilfeldige parametre:

$$(u_1, h_1, w_1), ..., (u_l, h_l, w_l) \in \mathbb{G}_p^3, \qquad g,\ v \in \mathbb{G}_p, \qquad g_q \in \mathbb{G}_q,$$

og en eksponent $\alpha \in \mathbb{Z}_p$. Alle disse parameterne beholdes som den skjulte nøkkelen SK [25]. Så velges $3l + 1$ tilfeldige faktorer fra G_q:

$$(R_{u,1}, R_{h,1}, R_{w,1},), ..., (R_{u,l}, R_{h,l}, R_{w,l},) \in \mathbb{G}_q \qquad \text{og} \qquad R_v \in \mathbb{G}_q$$

Til offentlig nøkkel, PK, brukes beskrivelsen av gruppen $\mathbb{G}$ og verdiene:

$$g_q, \quad V = vR_v, \quad A = e(g, v)^\alpha \ ,$$

$$\begin{pmatrix} U_1 = u_1 R_{u,1}, & H_1 = h_1 R_{h,1} & W_1 = w_1 R_{w,1} \\ & ... & \\ U_l = u_l R_{u,l}, & H_l = h_l R_{h,l} & W_l = w_l R_{w,l} \end{pmatrix}$$

Kryptér(PK, $\mathcal{I} \in \mathbb{Z}_m^l, M \in \mathcal{M} \subseteq \mathbb{G}_T$): Vi lar $\mathcal{I} = (\mathcal{I}_1, ..., \mathcal{I}_l) \in \mathbb{Z}_m^l$. Krypteringen foregår på følgende måte:

- Velg en tilfeldig $s \in \mathbb{Z}_n$ og en tilfeldig Z, $(Z_{1,1}, Z_{1,2}), ..., (Z_{l,1}, Z_{l,2}) \in \mathbb{G}_q$.
- Gi ut kryptert tekst:

$$C = \left(C' = MA^s, C_0 = V^s Z, \begin{pmatrix} C_{1,1} = (U_1^{\mathcal{I}_1}, H_1)^s Z_{1,1}, & C_{1,2} = W_1^s Z_{1,2} \\ \\ C_{l,1} = (U_l^{\mathcal{I}_l}, H_l)^s Z_{l,1}, & C_{l,2} = W_l^s Z_{l,2} \end{pmatrix} \right)$$

GenererSymbol(SK, $\mathcal{I}_* \in \sum_*^l$): der altså $\mathcal{I}_* \in \sum_*^l$ blir vår $\langle P \rangle$ som beskrevet for PECKS. $\langle P \rangle$ er da $\mathcal{I}_* = (\mathcal{I}_1, ..., \mathcal{I}_l) \in \{\mathbb{Z}_m \cup \{*\}\}^l$. Vi lar

S være mengden av alle indexer i slik at $\mathcal{I}_i \neq *$. For å generere et symbol for predikatet $P^{HVE}_{\mathcal{I}_*}$ velges tilfeldig $(r_{1,1}, r_{i,2}) \in \mathbb{Z}_p^2$ for alle $i \in S$ og gi ut :

$$TK = \left(\mathcal{I}_*, K_0 = g^\alpha \prod_{i \in S} (u_i^{\mathcal{I}_i} h_i)^{r_{i,1}} w_i^{r_{i,2}}, \forall i \in S : K_{i,1} = v^{r_i,1}, K_{i,2} = v^{r_i,2} \right)$$

Forespørsel(TK, C_t): beregner da

$$M \leftarrow C' / \left(e(C_0, K_0) / \prod_{i \in S} e(C_{i,1}, K_{i,1}) e(C_{i,2}, K_{i,2}) \right)$$

og gir enten ut meldingen (i kryptert form) eller ingenting.

6.4 Delegering, en utvidelse av HVE

En mulig utvidelse av HVE gir muligheten for at en brukergruppe skal kunne betro en undergruppe begrenset tilgang til data gjennom delegering [23, Kapittel 3]. Dersom en gruppe i politiet får tilgang til data om en bestemt person/bruker under et visst tidsrom, vil de kanskje gi tilgang til data fra en spesiell IP-adresse denne brukeren har anvendt, videre til en undergruppe i politiet som skal etterforske denne spesifikke saken. Da kan de *delegere* tilgang som er mer restriktiv enn den tilgangen de selv har. Delegering krever en utvidelse av HVE-systemet til systemet presentert i [23, Kapittel 3]. Delegering handler altså om å begrense antallet personer som får kunnskap om dataene til et minimum.

6.5 Multidimensjonelle spørringer på kryptert data

I [24] presenteres et kryptosystem som håndterer multi-dimensjonelle søk eller spørringer på kryptert data (MRQED). I artikkelen er problemstillingen noe lik problemstillingen for dette arbeidet, der det lagres revisjonsspor

fra datanettverket, men kun noen brukere i en gitt gruppe/organisasjon kan få tilgang til dataene gjennom en sentral autoritet. Når revisjonssporene tenkes å inneholde misstenkelig data, som spor av virus eller ormer, kan brukeren forespørre den sentrale autoriteten om tilgang til dataene. Dersom tilgang gis, skal brukeren kun få tilgang til spesifiserte data, og ikke hele databasen. For å forhindre misbruk er det ønskelig at brukeren får vite så lite som mulig fra søket etter dataene, og at det skal være nødvendig og dekryptere etterpå for tilstrekkelig innsyn i detaljene. Søket med MRQED skjer på forhåndsbestemte parametre som tid, t, adresse til datakilden, a, og portnummer, p. Dette er likt situasjonen der politiet kan forespørre tilgang til dataene de trenger til etterforskning gjennom tillatelse fra PT og rettsvesenet. Algoritmen som genererer nøkler til kryptering og dekryptering er gitt ved:

Setup$(\sum, \mathbb{L}_\delta)$: Tar inn en sikkerhetsparameter $\sum$ og et D-dimensjonellt gitter $\mathbb{L}_\delta$, og gir ut en offentlig nøkkel, PK, og en master skjult nøkkel, SK.

Så krypteres dataene med en krypteringsalgoritme gitt ved:

Kryptér(PK, $\mathbf{X}$, M): Tar inn den offentlige nøkkelen PK, et punkt $\mathbf{X}$, og en melding M fra en supermengde $\mathcal{M}$, og gir ut en kryptert tekst C_t.

Søkeresultatene må så dekrypteres med riktig nøkkel for å kunne forsyne politiet med ytterligere, og nødvendige, detaljer. I [24] må alle parameterne stemme overens med søkekriteriene. Dersom minst én av parameterne faller utenfor det spekteret det søkes på for denne parameteren, vil søket mislykkes. Denne egenskapen ved systemet spesifiseres ytterligere i algoritmen som skal beregne hvilken nøkkel det skal dekrypteres med. Hyper-spekterne i $t \in [t_1, t_2]$, $a \in [a_1, a_2]$ og $p \in [p_1, p_2]$ utformer et hyper-rektangel $\mathbb{B}$. I søket må altså et punkt $\mathbf{X}$ falle innenfor hyper-rektanglet $\mathbb{B}$. I beregningen for hvilken nøkkel som skal brukes til å dekryptere dataene brukes PK, SK og $\mathbb{B}$:

BeregnDekrypteringsnøkkel(SK, PK, $\mathbb{B}$): Tar inn den skjulte nøkke-

len SK, den offentlige nøkkelen PK og hyper-rektanglet $\mathbb{B}$, og gir ut en dekrypteringsnøkkel, DK, for $\mathbb{B}$.

Kun de meldingene der $\mathbf{X} \in \mathbb{B}$ kan nå utleveres og dekrypteres:

$$\textbf{\textit{Dekryptér}}(PK, DK, C_t) \begin{cases} M & \text{hvis } \mathbf{X} \in \mathbb{B} \\ \perp & \text{hvis } \mathbf{X} \notin \mathbb{B} \end{cases}$$

Dette kapittelet har presentert flere alternative kryptoalgoritmer for mulig bruk til kryptering og dekryptering, søk i krypterte data og aksesskontroll i en eventuell ny, teknisk lagringsløsning. Den minst avanserte løsningen var å utelate bruken av kryptoalgoritmer, og heller lagre dataene ukrypterte. Mekanismer for aksesskontroll ble da vesentlige for informasjonssikkerheten. Ved kryptering av lagret data var både IBE, PECKS og HVE, Delegering, og MRQED alternativer for å håndtere søk i dataene, og krypterings- og dekrypteringsmekanismer, men med ulikt fokus og ulike system- og situasjonskrav og hovedfunksjonaliteter.

Kapittel 7

Forslag til lagringsløsning

I dette kapittelet presenteres de administrative metodene som trengs for å håndtere en ny teknisk løsning, med rot i standardene ISO/IEC 27000-serien [17], [18] og ETSI [19]. En teknisk løsning som skal tilfredsstille den aktuelle situasjonen og de kravene som er satt presenteres også, med utgangspunkt i de samme standardene. Tilbydere av elektronisk kommunikasjon krypterer og lagrer data i en database. Når det er nødvendig kan politiet forespørre en høyere instans, rettsvesenet og PT, om tilgang til data som i klartekst tilfredsstiller visse verdier (tidsrom, telefonnummer, IP-adresse eller personnavn). Ved innvilgelse skal tilbyder søke opp dataene for politiet og politiet skal kun motta de dataene som forespørres. All annen informasjon skal holdes skjult og kryptert, også for tilbyder.

7.1 Administrativ løsning

Som nevnt har tilbydere av elektroniske kommunikasjonstjenester allerede erfaringer med å lagre data som ikke skal være tilgjengelig for eksterne brukere. Dette gjelder både data om kommunikasjonen til hver abonnent som brukes til fakturering, og også personalopplysninger eller forretningshemmeligheter de ikke vil at andre utenfor organisasjonen skal ha tilgang til. Ved lagring og hemmelighold av ulike typer data må tilbyderne være

forberedt på de risikoene slik informasjon fører med seg, og at risikoene kan modifiseres ved å implementere kontroll og administrativ håndtering innenfor informasjonssikkerhet.

Standarden ISO/IEC 27002:2005 presenterer informasjonsteknologi, sikkerhetsteknikk og administrasjon av informasjonssikkerhet [18]. Standarden "... fastlegger retningslinjer og generelle prinsipper for å opprette, iverksette, vedlikeholde og forbedre administrasjon av informasjonssikkerheten i en virksomhet." Målsetningene bak denne standarden er å gi praktiske retningslinjer for å oppnå allment akseptable mål for administrasjon av informasjonssikkerhet samt å skape tillit mellom samarbeidende virksomheter. Standardens målsetninger samsvarer med dette arbeidets mål om å skape tiltro til datalagring gjennom å prioritere informasjonssikkerhet i den foreslåtte lagringsløsningen, samt at det skal være tillit mellom brukerne av kommunikasjonstjenestene, tilbydere, politiet og påtalemyndigheter. ETSI-standarden presenterer detaljerte rutiner for bedre informasjonsikkerhet ved lagring av overvåkningsdata, hvilket passer den aktuelle situasjon med innsamling av elektroniske kommunikasjonsdata.

Det anbefales derfor som en del av de administrative rutinene ved en eventuell innføring av datalagringsdirektivet at standarder utviklet spesielt til håndtering av informasjonssikkerhet følges nøye. Dette inkluderer de nevnte standardene fra ISO/IEC og ETSI. Jo strengere og grundigere rutinene i disse standardene følges, jo bedre oversikt vil både tilbydere og eventuelle tredjeparts sikkerhetskontrollører få over sikkerheten og eventuelle brudd på sikkerheten.

7.2 Teknisk løsning

For å tilfredsstille alle kravene fra seksjon 5.1 og 5.2, trenger vi en teknisk løsning som er så sikker som mulig, der informasjonen som lagres krypteres, men enkelt søkes opp og dekrypteres dersom de riktige tillatelser for uthenting av data er gitt. I bunnen av den tekniske løsningen trenges det et søkbart

kryptosystem som benytter seg av asymmetriske nøkler. Dette arbeidets foreslåtte lagringsløsning er basert på PECKS [21] og HVE [22] presentert i seksjon 6.3.

7.2.1 Tilpasning av valgt kryptoalgoritme

For å tilpasse PECKS og HVE til den aktuelle lagringssituasjonen må det defineres hva verdiene i kryptoalgoritmene representerer og vi må se på de utfordringene denne aktuelle situasjonen skaper for disse generelle kryptoalgoritmene. Først defineres det hvilke verdier det kan søkes på, altså hva parameteren I skal være. M inneholder alle data for en gitt person med et gitt utstyr og tilhørende identifikasjoner som eventuelt blir påkrevd lagret av direktivet. Av alle disse lagrede dataene vil personnavn, telefonnummer (som A-, B-, eller C-nummer) eller IP-adresser, og tidsrom være essensielle søkefelter. Dette fordi det er ganske sannsynlig at det er en eller flere av disse parameterne politiet vet om fra før når de gjør etterforskning, og som de vil søke etter mer utfyllende data om. I defineres derfor til å utgjøre tuppelet $(p,\ n,\ a,\ t)$ der:

- p er personnavn;
- n er telefonnummer;
- a er IP-adresse;
- t er tidsrom;

Kun de meldinger der alle parameterne stemmer med søkekriteriene vil gi $P(I) = 1$. Dersom politiet ikke besitter verdier for alle parameterne, kan de bruke den definerte likegyldige verdien *. Hvis politiet for eksempel mottar informasjon om en IP-adresse som viser seg å være brukt til noe ulovlig og utlevering av data innvilges, kan politiet få tilbyder til å søke i databasen etter $(a \in [a_1, a_n] \cap t \in [t_1, t_n])$ og så sette $(p = *, n = *)$, dersom de mangler informasjon om navn eller telefonnummer.

HVE-modellen utfordres på noen punkter når den skal tilpasses situasjonen der lovpålagt datalagring er innført. I følge straffeprosesslovens §210 om utlevering av trafikkdata (ting) [7] og i samsvar med dagens praksis (seksjon 2.3.2) må politiet ha rettslig kjennelse dersom de skal få utlevert trafikkdata fra tilbyder. Politiet skal dermed ikke ha tilgang til annen data enn den de har fått tillatelse til. Dette oppsummeres slik:

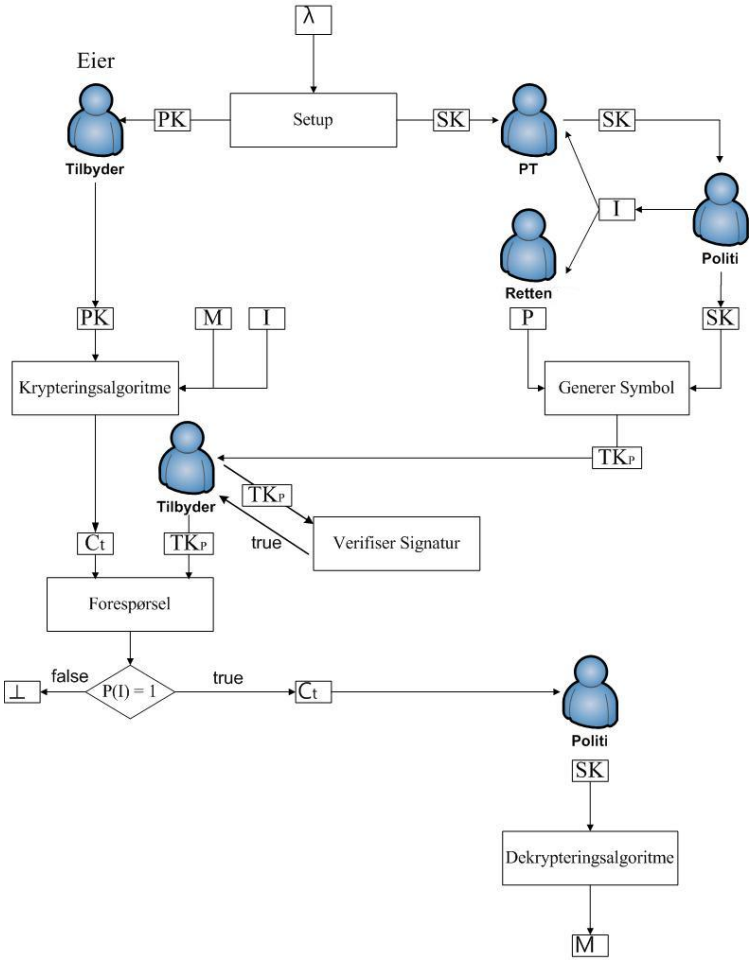
1. Politiet skal ikke ha mulighet til å lage et predikat og forespørsler som alltid gir ut true (sann), slik at alle data i databasen kan hentes ut på én forespørsel;

2. Politiet bør ikke ha mulighet til å generere symboler alene, da dette medfører at politiet i teorien kan lage forespørsler etter så mye data de ønsker;

3. Tilbyder skal kunne sjekke at politiets forespørsel er godkjent av retten;

For å kompensere for utfordring én, og at (verken PT, retten eller) politiet ikke skal kunne lage predikater eller forespørsler som alltid er sanne og dermed gi ut hele databasen med trafikkdata, kan hver forespørsel kun hente ut én kryptert melding. Én melding her er logger fra én person, én IP-adresse eller ett telefonnummer, under ett bestemt tidsrom. Altså resultatet av en spørring med ett tuppel, der maks tre av verdiene kan være likegyldige (*). Data om flere brukere bør ikke kunne hentes ut på samme forespørsel. For å kompensere for utfordring to, og at (verken PT, retten eller) politiet ikke skal kunne lage symboler alene, og dermed ikke generere så mange forespørsler de vil selv, fordeles informasjonen som trengs for å generere symboler mellom retten, PT og politiet. Kun retten kan generere predikater, mens politiet får tilgang til riktig skjult nøkkel fra PT som kan brukes til å generere symboler i samarbeid med retten, og dekryptere uthentet melding. For å kompensere for utfordring tre signerer retten predikatet og symbolet før spørringen sendes til tilbyder. Tilbyder må da verifisere at retten har signert predikatet og at politiet dermed har fått godkjent tilgang til en begrenset mengde data. Med disse nye kravene kan vi sette opp

hvordan flyten i lagring og uthenting, kryptering og dekryptering bør foregå (Figur 7.1).



Figur 7.1: Flytskjema for HVE-modell med entieter og forholdet mellom dem.

Hendelsesforløpet for bruk av lagret trafikkdata vil da gå som dette. Tilbyder mottar krypteringsnøkkel, og krypterer og lagrer de krypterte dataene i en database. Når politiet ser det nødvendig i forbindelse med etterforskning og bruke trafikkdata, sender de en forespørsel til PT og retten med beskrivelse av hva de har av informasjon (I) som kan brukes til å søke etter ønsket data (som er M). Dersom PT og retten samtykker, deler PT den riktige skjulte nøkkelen med politiet, og retten lager et predikat og signerer det. Et symbol blir generert ved hjelp av det signerte predikatet og den skjulte nøkkelen SK. Symbolet oversendes tilbyder som verifiserer at politiet har fått godkjent tilgang til en begrenset mangde data gjennom rettens signatur på predikatet. Kun dersom verifikasjonen bekreftes vil et søk i databasen utføres av tilbyder. Søkestrenger genereres, og det krypterte resultatet sendes tilbake til politiet. Under oversendelse av data kan det være lurt og legge til integritetssjekk på dataene, slik at politiet vet at dataene ikke har blitt tuklet med under overføring. Dataene kan beskyttes gjennom hashing-algoritmer ("hakke-algoritmer") eller hash-MAC (HMAC)-algoritmer. Slike integritetsmekanismer er ikke synlige i modellen. Politiet dekrypterer så mottatt data med sin skjulte nøkkel. Denne skjulte nøkkelen skal altså kunne brukes til å lage symbol, og dermed en forespørsel, samt til å dekryptere den krypterte meldingen som blir resultatet. Det er verdt og notere seg her at tilbyder sees på som eier av dataene, hvilket betyr at de har rådighet over dataene der de er lagret. De kan ikke generere forespørsler selv da de ikke vet hva predikatet eller den skjulte nøkkelen er, men de må likevel søke for politiet da vi antar at de har rådighet over selve databasen enten gjennom en lokal lagringsløsning eller en mellomløsning der flere tilbydere deler på lagringsmediet.

Distribusjon av nøkler

Maskinen som kjører *Setup*-algoritmen bør plasseres på et fysisk sikret sted, kanskje hos PT, som er den enheten som gir tillatelse til fritak fra taushetsplikt hos tilbydere gjennom å tillate politiet bruk av de lagrede dataene. Vi diskuterer ikke videre *hvordan* nøklene overføres til politiet, men sikre

kanaler kan være å foretrekke.

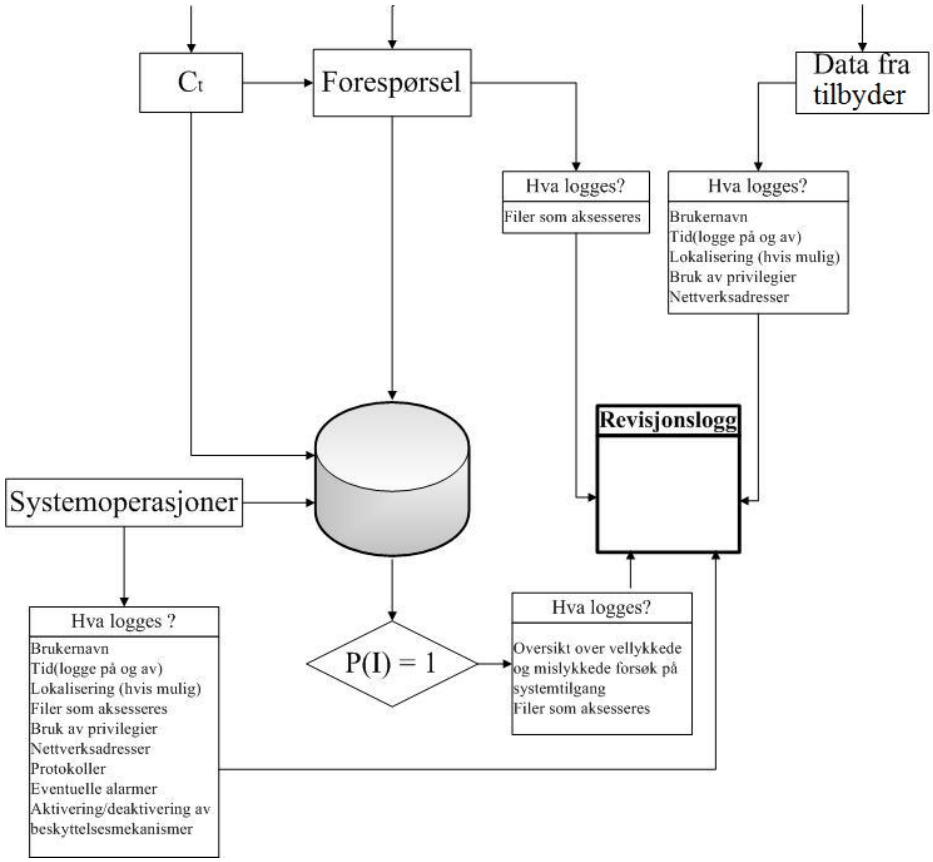
Lagring i database

Hver enkelte lagringsløsning kan representeres i modellen (Figur 7.1) for dette systemet. Enten flere tilbydere samarbeider om en felles løsning, eller om hver tilbyder har en egen løsning, kan de motta nøkkel for kryptering og lagre i sin database som i modellen. Dersom flere aktører deler database, mottar de alle (samme) nøkkel fra samme *Setup*-algoritme, krypterer dataene med den og lagrer dataene i likt format i tilhørende database. Uthentingsprosedyren er den samme uansett lagringsløsning. Dette vil ligne dagens praksis, da retten kan pålegge ulike tilbydere med ulike løsninger utlevering av data (seksjon 2.3.2).

Opprettelse av revisjonslogg

Alle forespørsler og resultater, brukernavn, tidsrom, nettverksadresser, bruk av privilegier, protokoller, alarmer som eventuelt utløses, aktiveringer eller deaktiveringer av beskyttelsesmekanismer lagres til en revisjonslogg (Figur 7.2) (jf. seksjon 5.1.1). Revisjonsloggen bør også opprettes for å imøtekomme de svakhetene som påpekes av ETSI [19] (seksjon 4.2) for å gi de som arbeider med systemet på noen som helst måte en ansvarsfølelse ovenfor sine handlinger. Det unike brukernavnet til brukere hos tilbyder sporer hvem som gjør forespørselen for bedre oversikt over bruken av databasen med trafikkdata. Brukernavn hos systemoperatører forteller hvem som har gjort hvilke operasjoner på systemet, og vil også gi bedre oversikt og sporbarhet over handlinger som inkluderer endringer i databasen. Revisjonsloggene som tilhører dataene det er gjort revisjoner på bør lagres like lenge som selve dataene, altså mellom 6 måneder og 2 år (jf. seksjon 2.2.1). Operasjoner på selve systemet som ikke gjelder uthenting av data bør også lagres i revisjonsloggen. Dette inkluderer vedlikehold av databasen, forebygging av feil, gjenopprettelse og oppdateringer av programvare. Hvilken type operasjon som ble gjort når og av hvem bør lagres for å kunne spore endringer. Hvordan autorisasjon til systemoperasjoner blir håndtert spesifiseres ikke

nærmere her.



Figur 7.2: Data som skal inkluderes i revisjonsloggen.

Sletting av data

I følge datalagringsdirektivet skal data lagres kun fra 6 måneder opptil 2 år, og ikke lenger. Dataene skal slettes etter lagringsperioden som hvert land fastsetter. Det bør altså forsikres av hensyn til personvernet at dataene blir

helt borte etter lagringsperioden, slik at de på ingen måte kan gjenopprettes. Data som muligens blir lagret på flere lokasjoner, må slettes fra alle fysiske lagringmedium, uavhengig av hvor eller hvordan de er lagret. I modellen kan man for eksempel ta utgangspunkt i hvilket tidspunkt hver datanotering lagres på, og så overskrive disse dataene med vilkårlig data når lagringstiden løper ut. Da forhindrer vi at dataene bare "slettes", men ikke ødelegges, og kan gjenopprettes. Overskriving av data som skal slettes anbefales også av ETSI [19].

Del III

Evaluering

Kapittel 8

Diskusjon

I dette kapittelet diskuteres fordeler og ulemper med den modellen som ble presentert i kapittel 7 og fordeler og ulemper med alternative modeller. Det fokuseres på fordeler og ulemper ved sikkerheten rundt de ulike lagringsløsningene, samt på kostnadsaspektene.

8.1 Evaluering av presentert løsning

Ingen løsninger er perfekte, og også løsningsmodellen presentert i dette arbeidet har svakheter. I denne seksjonen identifiseres noen av disse svakhetene, slik at de svake punktene enten senere kan erstattes av andre løsninger eller følges opp grundig. Svakheter ved både den administrative delen den tekniske kryptoalgoritmen identifiseres. Hvorfor HVE-kryptoalgoritmen likevel ble valgt i løsningen vises av fordelene denne kryptoalgoritmen har i modellen sammenlignet med de alternative kryptoalgoritmenes fordeler, som også diskuteres her.

8.1.1 Sikkerhetsaspekter ved den presenterte modellen

Denne seksjonen tar for seg sikkerhetsmessige svakheter eller ulemper i de administrative og tekniske løsningsrutinene som er presentert i del II: Kapit-

tel 4, 5 og 6.

Administrativ sikkerhet

I forslaget til en spesifikk, ny lagringsløsning, er høy sikkerhet prioritert fremfor lave kostnader, gjennom bruk av standardiserte løsninger og kjente kryptosystemer. I et system basert på omfattende standarder, der sertifiserbare løsninger gjør verifisering og oppfølging fra tredjeparter mulig, vil man lettere kunne identifisere brudd på sikkerheten både i det administrative og i det tekniske miljøet. Når lister over regler for personellsikkerhet, fysisk sikkerhet og prosedyrer (seksjon 5.1) er definerte og godt informert om til berørte personer, vil avvik fra rutinene kunne identifiseres lettere enn om det ikke fantes fastsatte rutiner og regler. Sikkerheten rundt dataene som lagres vil påvirkes av graden av grundighet rundt innføringen av nye administrative rutiner, samt forhåndsregler og forsiktighet som følges av alle personer som er i kontakt med systemet. Å følge grundige standarder for praktisering av informasjonssikkerhet vil være en fordel kontra å bruke egendefinerte retningslinjer for praktisering i hver organisasjon. Slike standarder eksisterer allerede og vi har henvist til disse under arbeidet med de administrative rutinene rundt en ny lagringsløsning. Standarder er standarder fordi de er nøye vurderte og godt dokumenterte, hvilket gir et godt grunnlag for retningslinjer, som lett kan følges opp. Man kan også følge opp alle tilbydere og entiteter i lagrinsløsningen på samme måte, og sammenligne grundigheten av praksisen hver tilbyder følger, uavhengig av type lagringsløsning. Dersom alle tilbydere velger sine egne retningslinjer og sine egne praktiseringsmetoder blir det vanskeligere å sammenligne hvor grundig informasjonssikkerhet prioriteres hos hver tilbyder. Hvilket organ som eventuelt skal gjøre en slik oppfølging hos tilbyderne, og hvilke konsekvenser det vil få for tilbyder dersom de standardiserte kravene ikke følges, diskuteres ikke videre her.

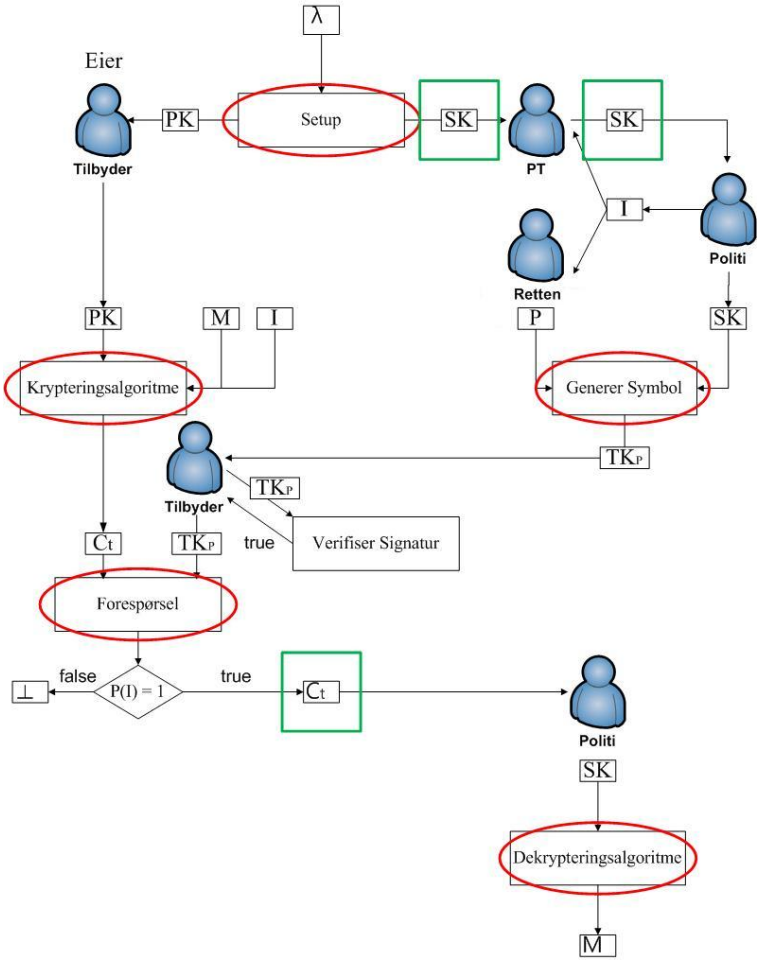
Teknisk sikkerhet

Sikkerheten i den presenterte tekniske modellen er kun så sterk som sitt svakeste ledd. For å vurdere sikkerheten i modellen må leddene der tekniske

sikkerhetsbrudd kan forekomme identifiseres. Sikkerheten i modellen baseres på følgende punkter:

- Den underliggende sikkerheten i de kryptografiske funksjonene (algoritmene);
- Sikkerheten under overføring av data mellom entiteter;
- Sikkerheten rundt selve lagringen av de aktuelle dataene;
- Grundigheten rundt prosessen ved å hente ut data;
- Brukernes forholdsregler og forsiktighet rundt systemets prosesser;

Den underliggende sikkerheten i de kryptografiske funksjonene (algoritmene) (Figur 8.1) avhenger av sikkerheten i HVE. Sikkerheten i HVE-kryptoalgoritmen som helhet defineres ved hjelp av den sammensatte 3-delte Diffie-Hellman antakelsen (composite 3-party Diffie-Hellman assumption) (C3DH) [25]. Antakelsen er basert på intuisjonen om at det er vanskelig å teste for Diffie-Hellman tupler i en undergruppe av orden p hvis elementene som skal testes har en tilfeldig komponent i en undergruppe av orden q (se vedlegg eller [25] for utredning). Dersom det antas at den sammensatte 3-delte Diffie-Hellman antakelsen holder, kan ingen motstander eller fiende bryte systemet innen polynomisk tid med en betydelig fordel. Når HVE brukes, skjules også attributtene i I fra alle utenforstående. Kun politiet, rettsvesenet og PT vet om I. En potensiell motpart fra utsiden vil derfor ikke se hvilke attributter det er snakk om at spørringen og påfølgende melding inneholder, kun at spørringen stemmer. Dette gir en bedre sikkerhet enn med andre kryptoalgoritmer, inkludert Identity Based Encryption (IBE) [26], [27] og Multi-Dimentional Range Query over Encrypted Data (MRQED) [24] (kapittel 6), som også tilbyr søk på krypterte data. Det avsløres ikke hvilke av attributtene som har bestemte verdier og hvilke som har tilfeldige verdier (*) og ikke hva de bestemte attributtene er. Hvilke av attributtene M inneholder vil altså ikke avsløres.



Figur 8.1: Potensielle svake tekniske ledd i den presenterte modellen. Ellipsene identifiserer algoritmene, kvadrater identifiserer overføring av sensitiv data.

Sikkerheten under overføring av data mellom entiteter (Figur 8.1) forsterkes av at hver overføring ikke inneholder tilstrekkelig informasjon til og verken søke i eller hente ut og dekryptere data som er lagret. Den offentlige nøkkelen, PK, er som sagt *offentlig* og kan potensielt anskaffes av hvem som helst. Får man tak i denne kan man kun kryptere data, og ikke dekryptere. Kun tilbyder har tillatelse til å søke i dataene, så dersom en trusselaktør skulle fått tak i både den skjulte nøkkelen og et predikat, *og* fått vite hvordan man genererer symboler, måtte trusselaktøren fortsatt maskert seg som tilbyder og fått tilbyders rettigheter for å kunne søke etter data og dekryptere resultatet. Å skaffe kun symbolet og tilbyders rettigheter til søk vil ikke ha betydning for selve dekrypteringen, og eventuell uthentet informasjon vil fortsatt være kryptert. Dersom en trusselaktør får tak i den skjulte nøkkelen, SK, og de dataene som overføres til politiet etter et vellykket søk, kan trusselaktøren potensielt dekryptere disse og lese dem i klartekst. Selv om dette begrenser tilgangen til dataene til resultatene av rettslig tillatte søk, er det fortsatt en svakhet at andre enn politiet skal kunne lese dataene. Det anbefales derfor at sikre kanaler brukes både når SK og data, som resultat av vellykkede søk, overføres til politiet.

Sikkerheten rundt selve lagringen refererer til om dataene som er lagret ligger i klartekst eller som kryptogram, og tilgangsmetodene til disse. Tilgangsmetodene er allerede diskutert over, og om det er nødvendig å kryptere dataene eller om tilgangsmetodene gir sterk nok sikkerhet til at de kan lagres i klartekst utdypes under seksjon 8.1.4. Det er ingen tvil om at det gir en bedre sikkerhet for lagringen dersom dataene er krypterte. Dersom dataene er krypterte kan ingen som tilfeldigvis kommer over data, for eksempel ved vedlikeholds- eller andre systemoperasjoner, lese hva som står der, og heller ikke lete blant dataene, siden det ikke er synlig hva som er hva og hvordan dataene ligger sortert. Om kryptering er nødvendig tatt i betraktning hvilken type informasjon dataene som lagres inneholder diskuteres ikke videre her.

Grundigheten rundt prosessen ved å hente ut data gjenspeiles av de administrative prosessene hos de ulike entitetene. Hvor hardt innarbeidet ru-

tinene rundt datalagring, uthenting, og bruk av informasjon er, har noe å si for sikkerheten rundt lagringen. Graden av grundighet vises altså gjennom den administrative sikkerheten, nevnt over (seksjon 8.1.1). Det samme gjelder brukernes forholdsregler og forsiktighet rundt systemets prosesser. Den valgte tekniske sikkerheten er minst like god som, om ikke bedre enn, de presenterte alternative løsningene. Man kan aldri være sikker på at det ikke finnes smutthull i den tekniske løsningen, og 100% sikker vil en teknisk løsning kanskje aldri kunne bli. Men ved å distribuere både ansvarsområder og kunnskap kan sannsynligheten for ulovlig bruk av løsningen av utro tjenere hos tilbyder, politiet, retten eller PT reduseres.

8.1.2 Kostnadsaspekter ved den presenterte modellen

En ulempe med et både teknisk og administrativt omfattende system, er kostnadene utvikling, tilpasning og ny kunnskap, bruk og vedlikehold fører med seg. Den faktoren som påvirker kostnadene mest i Teleplans økonomiske analyse av datalagringsdirektivet er utviklingstiden av det nye lagringssystemet [6]. I [6, Kapittel 5.9] presenteres resultatene av analysen og det Teleplan kommer frem til er: "Økes utviklingstiden fire ganger, øker kostnadene med 35-50 prosent avhengig av volum og lagringsløsning. Hvis datavolumet økes fire ganger fra 3TB til 12TB økes kostnadene med 1- 4 prosent avhengig av lagringsløsning. Selv en økning fra 20TB - 80TB vil ikke gi en så stor kostnadsøkning som for en tilsvarende økning i utviklingstid." Selv om utviklingstiden varierer med mengden data og type lagringsløsning, enten den blir sentral eller lokal, vil utviklingstiden være den mest utslagsgivende faktoren. Jo mer omfattende systemet som skal implementeres er, som påvirkes av hvor omfattende de administrative og tekniske delene er, jo mer kostbart vil det bli å utføre lagringen. Å innføre en database der de lagrede dataene krypteres kontra å bruke en database med data lagret i klartekst vil dermed gjøre implementeringen dyrere [8], [19]. Dette betyr at utviklingen og innføringen av den foreslåtte modellen sannsynlig vis vil koste betydelig mer enn utvikling og innføring av en modell med data lagret i klartekst. Stipendiat ved institutt for privatrett, senter for rettsinformatikk ved Universitetet i Oslo (UiO), Tobias Mahler, påpeker også under et åpent

lunsjforedrag de ekstra kostnadene en asymmetrisk kryptert løsning vil føre med seg, gjengitt i Dag-Rune Vollens artikkel i Computerworld om datalagringsdirektivets innføringskostnader [28]. Mahler påpeker at kravene den tyske Forfatningsdomstolen har satt til lagringen var svært strenge, og disse kravene ligner dette arbeidets presenterte krav til implementering av en ny lagringsløsning. Dette inkluderer separat lagring av data i lager som ikke er tilknyttet offentlig nettverk og asymmetrisk kryptering av data, og Stopp DLDs styremeldem Torgeir Waterhouse mener slike tiltak kan mangedoble kostnadene til datalgring [28]. Det er derfor ingen tvil om at den løsningen som presenteres i dette arbeidet ikke er det billigste alternativet ved en innføring av datalagringsdirektivet i Norge. Antall arbeidstimer som også legges inn i den administrative delen vil kunne regnes som utviklingskostnader, og vil være høyere jo mer omfattende rutineendringene rundt det nye systemet er. Å følge de nevnte omfattende standardene ned til minste detalj, vil kunne gi høyere kostnader enn dersom dagens rutiner beholdes uforandret. Men har mange bedrifter både teknologi og gode administrative retningslinjer fra før vil timeverket reduseres, og kostnadene reduseres. Hvor godt hver tilbyder er skodd fra før vil variere fra tilbyder til tilbyder, og kostnadene kan derfor også variere.

Å maksimere sikkerheten er prioritert i dette arbeidet fremfor å minimere utviklingskostnadene fordi det har vært så mye debatt om datalagringsdirektivet og hvordan en eventuell innføring av direktivet vil påvirke personvern. Det vil også være urimelig og ikke prioritere sikkerhet når det er snakk om så mye data lagret om hver person over lengre tid. En samensatt og kompleks løsning med en kryptert database, lik den som presenteres, vil muligens forverre proporsjonene mellom kostnad og nytteverdi, men forbedre proporsjonene mellom nytteverdi og personvern. Om fordelene eller ulempene ved disse proporsjonene vektlegges høyest kommer an på hvilke faktorer man velger å prioritere. Å beregne eksakte tall for hva utviklingen av den presenterte løsningsmodellen vil koste er utenfor dette arbeidets omfang. Det spekuleres heller ikke i om de pengene som må brukes til utvikling av en modell lik den som presenteres kunne vært brukt anderledes til samme formål som presenteres i direktivet. Vi mener kun at dersom en ny lagringsløs-

ning må utvikles og innføres, vil vi prioritere sikkerhet gjennom grundige administrative rutiner med bakgrunn i presenterte standarder og en sammensatt og kompleks teknisk løsning med en kryptert database.

8.1.3 Presentert modell og personvern

De tekniske aspektene ved løsningen som foreslås, forsikrer at vanlig ansatte hos tilbyder eller operatør ikke har tilgang til trafikkdataene som lagres (seksjon 7.2.1). Vanlig ansatte har ikke tilgang da systemet er basert på offentlig-nøkkel kryptografi, og dataene krypteres med politiets offentlige nøkkel, PK, men dekrypteres med en annen, SK. Bruken av en kryptert database til lagring av trafikkdata forhindrer truslene mot personvernet i form av at ansatte hos tilbyder kan lese data og bruke dem til egen vinning [8]. Altså styrker en kryptert database, med offentlig-nøkkel kryptografi, lagringens forhold til personvern. Forslaget til en lagringsløsning forsikrer også at dataene er kryptert under overføring til politiet, og at databasens resterende data fortsatt vil være lagret hos tilbyder. Dette betyr at selv om politiet har mottatt data, kan de ikke søke etter andre data, ei heller oppdrive noen detaljer om annen data.

8.1.4 Utfordringer med den presenterte modellen

Denne seksjonen presenterer noen av de utfordringene den presenterte modellen må imøtekomme gjennom ulike eksterne og interne tiltak. Foreløpige mangler eller svakheter identifiseres og diskuteres.

Simulering

Dette arbeidets presenterte forslag til en ny lagringsløsning for den aktuelle situasjonen med lovpålagt datalagring er ikke simulert eller testet. Dette betyr ikke at modellen ikke fungerer for det gitte scenariet, men ved å simulere et reelt hendelsesforløp, med generering av nøkler, kryptering og dekryptering av data, forespørsel om tilgang, utveksling av informasjon, og med reelle trusler; kan svakheter i løsningen identifiseres bedre og flyten i saksgangen

samt eventuelle flaskehalser ville vist seg. Dette kunne gitt en bedre modell med teknologiske valg begrunnet i fysiske målinger. Dette kunne gitt et sterkere resultat.

Kompleks modell

Modellen som presenteres (seksjon 7.2.1) er relativt kompleks og sammensatt. Modellen inkluderer samarbeid mellom flere instanser, en kryptert database, flere kritiske algoritmer, kritiske parametre (SK og TK) og overføring av disse mellom entiteter. En sammensatt søke- og uthentingsprosess for lagrede data kan være bra, da det gjør det vanskelig for en trusselaktør å samle opp all informasjon som trengs for å gjøre et vellykket søk i databasen med resultater i klartekst. Men modellen kan også være unødvendig kompleks. Det er ikke sikkert det er nødvendig å kryptere dataene som lagres i databasen. Bruken av en ukryptert database gir en helt klart en mye enklere løsning, og mest sannsynlig en mindre kostbar løsning å utvikle. Problemet med en ukryptert database er at dersom en trusselaktør klarer å maskere seg som en bruker med administratorrettigheter innenfor tilbydernes brukergruppe, vil det være fri tilgang til hele databasen. I den presenterte modellen (seksjon 7.2.1) gir ikke en gang en forespørsel tilgang til hele databasen, da kun én melding, M, kan gis ut om gangen, og hver nye forespørsel krever et nytt predikat fra retten og et nytt symbol for spørringen. ETSI anbefaler ikke å lagre data i forbindelse med "overvåkning" i klartekst. De anbefaler at data er kryptert under hele lagringsperioden av sikkerhetsmessige hensyn [19, Avsnitt 7.6.1]. Siden dette arbeidets hovedfokus er informasjonssikkerhet, velges derfor en mer kompleks modell med færre og vanskeligere aksessmuligheter fremfor en enklere, sikkert billigere, løsning der sikkerheten er lavere. Fordi det også er så mye debatt rundt direktivet og dets innvirkning på personvernet, kan det være nødvendig og implementere en sammensatt og bedre sikret løsning med krypterte data i databasen. Dessuten kan det være enklere å starte med en kompleks lagringsløsning for så å eliminere de sikkerhetsaspektene man anser som

overflødige, enn å starte med en veldig enkel løsning og så bygge på den med sikkerhetsaspekter som skal fungere sammen.

Innsideangrep

Modellen er ikke sikker mot korrupthet eller innsideangrep. Dersom brukere i rettssystemet, PT og politiet sammen bestemmer seg for å finne info til egen vinning kan de sammen generere en forespørsel, med signert predikat av retten, og politiet kan sende denne til tilbyder. Tilbyder vil se denne forespørselen som en hvilken som helst annen forespørsel, søke i databasen og overlevere informasjonen. Denne typen trusler er forsøkt redusert ved å la brukere innenfor disse tre instansene ha ulike rettigheter og oppgaver i forhold til uthenting av informasjon i systemet. I og med at det er tilbyder som har rådighet over dataene og databasen, kan ikke korrupte brukere innenfor de tre andre instansene verken søke fritt i databasen eller hente ut informasjon om mange brukere på en gang. Men de kan forespørre data uten grunnlag dersom uhederlige brukere innenfor de tre samarbeider for å generere riktig forespørsel med riktig signatur. Sannsynligheten for at brukere innenfor politiet, PT og rettsvesenet kan være korrupte og samarbeider om å hente ut informasjon på en ulovlig måte er tilstede, men trolig ganske liten. Noen brukere bør uansett kunne ha tilgang til systemet for veldikehold eller gjenoppretting etter feil eller i nødssituasjoner. Selv om mye oppdatering, gjenoppretting og sikkerhetskopiering kan skje helautomatisk, er man i stort sett alle systemer avhengig av at noen brukere har tilgang som administrator til systemet. Man kan aldri garantere for at denne administratoren ikke er korrupt eller opptrer som en trusselaktør. Dette er vel også en av grunnene til at direktivet er heftig debattert: at man aldri kan være 100% sikker på at alle som jobber med systemet på noe vis, er hederlige mennesker og ikke korrupte, selv om sannsynligheten er liten. Det sikres heller ikke i modellen at tilbyder ikke kan lagre dataene sine et annet sted for egen tilgang. All data som genereres må derfor krypteres og lagres automatisk i databasen, men det kan fortsatt være muligheter for

tilbyder å samle opp informasjon om enkelte abonnenter. De bruker riktig nok deler av informasjonen allerede i dag til fakturering, men disse dataene skal i utgangspunktet slettes etter de har blitt brukt til sitt formål. Derfor skilles det også mellom data brukt til fakturering og data lagret for bruk i etterforskning, som også vil lagres noe lengre enn data gjør i dag. Svein Willassen påpeker også i sin rapport at en fordel i forhold til risikofaktorer for personvern kan være at man ved systemets implementasjonstidspunkt, bestemmer hvilken nøkkel det skal være mulig å hente ut data for senere [8]. Ved at systemet har slike egenskaper forhindrer man at man glir over i en situasjon der man til stadighet kan hente ut data på nye måter fordi dataene uansett ligger lagret. Om det blir mulig å hente ut eller lese data på stadige nye måter bestemmes av mekanismen for å hente ut data, det vil si generere symboler/forespørsler; og mekanismen for å dekyrptere de dataene man får ut. Når det gjelder endringer i prosessen for å hente ut data kommer dette an på hvordan predikatet, P, kan varieres til å generere symbolet TK_p sammenlignet med hvordan SK brukes i den samme genereringsprosessen. P i vårt tilfelle er $\mathcal{I}_* \in \sum_*^l$, og for å generere et symbol for predikatet $P_{\mathcal{I}_*}^{HVE}$ velges tilfeldig $(r_{1,1}, r_{i,2}) \in \mathbb{Z}_p^2$ for alle $i \in S$ (se seksjon 6.3). I og med at man velger tilfeldig $(r_{1,1}, r_{i,2}) \in \mathbb{Z}_p^2$ for alle $i \in S$, og ikke kan "bestemme" noe særlig over denne funksjonen for å endre på det vi vil ha ut, vil ikke predikatet kunne endres for å hente ut data på nye måter i fremtiden. Når det gjelder endringer i prosessen for å dekryptere data er denne den samme som den bestemt på implementeringstidspunktet. For dekryptering i fremtiden kan man kun bruke den skjulte nøklenen, SK, som det ble bestemt ved implementasjonstidspunkt skal kunne brukes til dekryptering. Det vil si at det ikke kan oppstå noen nye måter å dekryptere dataene på ved senere tidspunkt. Kun politiet kan dekryptere dataene med sin private, skjulte nøkkel, hvilket forhindrer andre instanser fra å kunne dekryptere data.

Angrep direkte på systemet

Fordi lagringssystemet lagrer de typer data det gjør, kan selve systemet bli et mål for kriminelle som enten vil blokkere systemets funksjoner eller tilegne seg lagret informasjon til egen vinning. Blokkering av systemets funksjoner kan foregå gjennom tjenestenektsangrep eller Denial of Service (DoS) angrep. Det er derfor ønskelig å beskytte infrastrukturen i kjernen av systemet på best mulig måte, og at man oppdager og forkaster slike typer angrep i systemets ytterpunkter. Dropping av angrepstrafikk kan gjøre gjennom lister for aksesskontroll, "blackholing" (som betyr å droppe all trafikk fra en bestemt IP-adresse eller domene), eller "null routing" (som betyr å droppe all trafikk som ikke har noen destinasjonsadresse) [19]. For at systemet skal fungere best mulig er det viktig at nedetiden til systemet er lavest mulig. For å begrense nedetiden er det viktig at sikkerhetsmekanismene rundt systemet er så bra som mulig, slik det er med alle datasystemer i dag. Det er forventet at lagringsløsningen, uavhengig av type løsning (lokal, sentral eller mellomløsning), forholder seg til kravene spesifiert under krav til prosedyrer (seksjon 5.1.1). Systemet må ha et moderne operativsystem, og applikasjoner som detekterer og beskytter systemet mot ondsinnede programmer, intregning og andre trusler, skal være på et høyt teknisk nivå. Selv om denne typen sikkerhet rundt systemet ikke beskrives i modellen, antas det at systemet beskyttes på best mulig måte på lik linje med andre lagringssystemer som lagrer sensitiv informasjon over lengre tid.

Sletting av data

Det er helt nødvendig at data som ikke skal lagres lenger slettes slik at de ikke kan gjenopprettes. Det anbefales i dette arbeidet at dataene overskrives, men det er ikke satt noen endelig mekanisme for dette. Ved å sette et tidsstempel på dataene som lagres, eventuelt ved bruk av tiden t som er en parameter i det søkbare feltet I, kan man implementere en automatisk sletting ved overskriving ved riktig tidspunkt. Dersom dataene blir hentet

ut i forbindelse med en etterforskning må (bør) de også lagres et annet sted for å unngå sletting mens de er i bruk, ellers må tidsstempelet endres.

Den presenterte løsningen for lagring av elektroniske kommunikasjonsdata har ikke blitt simulert gjennom dette arbeidet, og modellen er kompleks og sammensatt samtidig som den ikke er sikker mot innsideangrep. Systemet kan bli utsatt for angrep direkte, hvilket må motvirkes gjennom eksterne mekanismer, samtidig som interne mekanismer for sletting og overskriving av data må (bør) implementeres.

8.2 Evaluering av alternative modeller

I denne seksjonen diskuteres hvilke fordeler den presenterte løsningen har sammenlignet med om noen av de alternative presenterte kryptoalgoritmene (kapittel 6) hadde blitt brukt i løsningen. Dette for å fremheve hvorfor HVE ble sett på som den mest egnede kryptoalgoritmen for håndtering av kryptering og dekryptering av data, samt aksesskontroll ved et nytt datalagringssystem.

8.2.1 Sikkerhetsaspekter ved alternative modeller

Denne underseksjonen diskuterer hvorfor PECKS og HVE ble valgt fremfor de alternative presenterte kryptoalgoritmene med tanke på sikkerheten, og hvorfor den presenterte modellen med HVE besvarer problemstillingen og formålet med arbeidet på en best mulig måte.

Ukryptert database

Som nevnt (seksjon 6.1) var det et mulig alternativ og ikke bruke en kryptert database som lagringsløsning, men lagre i klartekst og bruke tilgangskontroll som hovedpunkt for sikkerheten. Dette var helt klart en mulig løsning, og kunne vært tilstrekkelig dersom vurderingen av graden av nødvendig skjerming av dataene tilsa dette. En løsning som involverte en ukryptert database ville helt klart vært mindre komplisert å opprette, og søk i dataene

kunne mulig vært gjort enklere. Selv om databasen var ukryptert kunne man fortsatt kryptert overføringen av dataene mellom tilbyder og politiet etter søk i databasen for å unngå tyvlytting og angrep gjort av mellommenn. Hvilken type kryptoalgoritme, symmetrisk eller asymmetrisk, som ville fungert best til dette formålet og hvilke ulike algoritmer innenfor valgt kryptoalgoritme, diskuteres ikke videre her. Dette ville ikke vært den optimale løsningen. For det første anbefaler ETSI at man krypterer all trafikkdata relatert til overvåkning og digital etterforskning under hele lagringstiden [19]. Dette både for bedre aksesskontroll og sikring av konfidensialitet og integritet. Dette er to viktige aspekter når vi skal lagre så mye detaljert data om privatpersoner. For det andre vil en ukryptert database påføre personvernet store trussler fordi ansatte hos tilbyder så "lett" kan aksessere databasen fra deres side. Ansatte hos tilbyder vil da potensielt ha kontinuerlig tilgang til hele databasen. Dette er en meget uheldig situasjon som vi vil forebygge ved å bruke en kryptert database. De ekstra kostnadene som påføres tilbydere ved å bruke en kryptert database har betydning i mindre grad enn de truslene en ukryptert database vil ha på personvernet.

Identitetsbaserte kryptosystemer

I motsetning til bruk av det presenterte forslaget til kryptosystem, måtte man ved bruk av IBE ha visst hvem som skulle lest meldingen før den ble kryptert. Siden meldingene (i det aktuelle tilfellet, trafikkdataene) skulle vært personlige med IBE, ville en slik praksis satt en grense for hvem som kunne ha dekryptert dataene i systemet. Uten å forutsi hvilken person som skal lese ut hvilke data, ville det vært vanskelig å bruke denne typen kryptoalgoritme til lagring. Dessuten var det ikke ønskelig at en bruker skulle kunne dekryptere mye data av gangen. "Politiet" kunne selvfølgelig vært satt som én bruker, eller man kunne opprettet flere generelle brukere for politiet, men de ønskelige søke- og tilgangsegenskapene ville likevel ikke blitt oppnådd med IBE. IBE var i likhet med HVE-systemet ikke sikkert mot korrupthet eller innsideangrep. Selv om det med IBE ville vært færre algoritmer som utgjorde potensielle svake ledd, ville det samtidig vært lettere å søke i databasen dersom man fikk tak i en bruker Bs skjulte/private nøkkel,

sammenlignet med om man fikk tak i den skjulte nøkkelen i det foreslåtte HVE-systemet. Fordi, med IBE kunne man dekryptert alle data som var kryptert med tanke på denne brukerens identitet, mens man med HVE kun kunne dekryptert de resultatene som eventuelt ble oversendt politiet dersom man skulle fått tak i disse under en oversending. Man kunne ikke ha søkt etter data i det presenterte systemet ved og kun besitte den skjulte nøkkelen. IBE var altså upraktisk i forhold til hvem som skulle kunne dekryptere hva i det sktuelle scenariet. Når det gjaldt korrupte brukere hos tilbyder kunne tilbyder, på samme måte som med HVE, gjemt unna data til eget bruk. Dette ville utgjort en trussel. Men tilbyder kunne ikke ha søkt fritt i databasen med IBE fordi IBE også er en asymmetrisk krypoalgoritme. På denne måten er IBE like sikker mot søk gjort av en tilbyder-bruker som HVE.

Multi-dimensjonelle spørringer på kryptert data

Som nevnt i seksjon 6.5 måtte man med MRQED visst alle parameterne man skulle søke etter, og alle måtte ha tilfredsstilt spørringen. Det hadde utgjort en ulempe i det sktuelle scenariet at alle parameterne i hyper-spekteret $\mathbf{X}$ hadde måttet ligget innenfor hyper-rektangelet $\mathbb{B}$. Det er ikke sikkert politiet hadde visst både navn, IP-adresse, telefonnummer og tiden de søkte etter. Som vist i brukerscenariene (seksjon 4.3) kunne det hendt politiet ikke var i besittelse av alle de nødvendige parameterne man kunne brukt som søkeparametre. Kanskje politiet hadde mottatt IP-adresser fra politidistrikt i andre land, der kun IP-adresse eller navn på en person var gitt. Dersom etterforskningen var god nok og saken kvalifiserte til et søk i databasen over lagret data, burde det vært mulig å søke i databasen uten å ha kjent alle parameterne. Forskjellen mellom den presenterte modellen med HVE, og MRQED, var at med HVE kunne man brukt tilfeldige verdier, *, dersom ikke alle parameterne var kjent eller lå innenfor et kjent spekter. Så her har bruken av HVE i modellen en fordel sammenlignet med om MRQED hadde blitt brukt.

Det er også to andre vesentlige hovedforskjeller på HVE og MRQED: sikkerhet; og størrelsesorden på offentlig nøkkel, krypteringstid og kryptert tekst. Det bevises ingen bedre sikkerhet for MRQED enn for HVE. Tvert imot konstaterer Shi et al. at sikkerhetsnotasjonen som kan bevises for HVE er sterkere enn for det systemet de selv presenterer, MRQED. I HVE [25] gjemmes attributtverdiene, I, til og med når meldingen er dekryptert. I MRQED avsløres attributtene når meldingen suksessfullt dekrypteres. HVE mer kostbar enn MRQED når det gjelder størrelsen på den offentlige nøkkelen PK, krypteringstid og størrelse på den krypterte teksten. Dersom D defineres til å være antall dimensjoner i attributtet og T til å være antall mulige diskrete verdier for hvert attributt, vil $O(DT)$ være størrelsesordenen på PK, krypteringstid og den krypterte teksten. MRQED oppnår en størrelsesorden på $O(D\log T)$ på de samme attributtene, og er derfor mer passende når D er liten og T er stor. HVE har med andre ord bedre sikkerhet gjennom skjult I, men dårligere størrelsesorden på noen attributter og parametre. Hva man velger å prioritere av dette bør være en vurderingssak for hvert scenario. Fordi det var andre unike egenskaper ved HVE som passet bedre til det aktuelle scenariet, ble HVE valgt til løsningen i steden for MRQED, selv om det strider med kravet om at uthenting av info skal være så effektivt som mulig (seksjon 5.2). En modell med MRQED ville heller ikke vært sikker mot innsideangrep. Dersom MRQED hadde blitt tilpasset det aktuelle scenariet, ville tilbyder mest sannsynlig vært den som utformet $\mathbb{B}$ for politiet, som da kunne dekryptert de dataene som passet akkurat til deres forespørsel $t \in [t_1, t_2]$, $a \in [a_1, a_2]$ og $p \in [p_1, p_2]$. Dersom en bruker med de riktige rettighetene innenfor entiteten tilbyder hadde vært korrupt, og bestemt seg for å finne data til egen vinning, hadde det ikke vært noen bestemt mekanisme for å hindre dette. MRQED er derfor i utgangspunktet like usikker som HVE når det kommer til innsideangrep. Selv om innsideangrep ikke er sannsynlig i stor skala, må mulighetene for at det skjer konstateres og identifiseres. Selv om politiet også med MRQED hadde trengt tillatelse fra både PT og rettsvesen for å sende forespørsel til tilbyder, er det ingen teknisk mekanisme som hindrer politiet og tilbyder fra å samarbeide om å hente ut data. Med HVE ble det satt opp at retten

og politiet må generere et symbol sammen for å kunne forespørre data fordi både predikat og nøkkel trengs for å lage et symbol. I MRQED ikke er en slik mekanisme. MRQED inkluderer bare algoritmene *Setup*, *Kryptér*, *Beregn-Dekrypteringsnøkkel* og *Dekryptér*. Et symbol lagd av predikater er ikke nødvendig. Dermed er det færre innsanser som må inkluderes i et innsideangrep i MRQED (politiet og tilbyder) enn med bruk av HVE (politiet, tilbyder, PT og retten). At flere brukere fra flere instanser må inkluderes kan gi en mindre sannsynlighet for slike innsideangrep. Notér også her at tilbyder kan potensielt gjemme unna data til eget bruk slik de også kan med alle de andre kryptoalgoritmene, og som de også kan med det systemet og de reglene som praktiseres i dag.

Delegering, en utvidelse av HVE

Sikkerheten i kryptosystemet med utvidet bruk av HVE var definert til samme nivå som sikkerheten i vanlig HVE [25]. Derfor ville ikke delegering gjort lagringssystemet mindre sikkert, sett bort fra at enda et ledd til med informasjonsoverføring hadde blitt innført, noe som kunne utgjort en sikkerhetstrussel. Grunnen til at muligheten for å delegere videre en mer restriktiv tilgang til data ikke ble presentert i dette arbeidets originale forslag til en lagringsløsning, er at det ikke ville vært *nødvendig* for en bedre sikkerhet rundt den lagrede informasjonen. Derfor ble denne funksjonen utelatt for å gjøre modellen så enkel og kostnadsbesparende som mulig ut fra de kravene som ble stilt til sikkerhet og kostnader.

8.2.2 Kostnadsaspekter ved alternative modeller

Denne seksjonen tar for seg kostnadsaspektene ved PECKS og HVE sammelignet med de alternative kryptoalgoritmene. Som nevnt over (seksjon 8.1.2) var utviklingstiden den faktoren som påvirkert kostnadene mest ved en ny lagringsløsning. Det blir ikke gjort noen videre kostnadsanalyse her, men det antas at kryptosystemer som er ca like avanserte vil ha ca like kostnader når det gjelder utvikling. Dette gjelder både foreslått løsning med HVE, IBE- og MRQED-kryptosystemer, samt den mulige utvidelsen

av HVE-kryptosystemer med delegering. De krevde alle at databasen var kryptert, og de benyttet ulike metoder for aksess til og søk i den krypterte databasen. Det alternativet som skiller seg ut er å la være å bruke en kryptert database, men heller bruke teknisk og fysisk tilgangsbegrensning som hovedmoment for informasjonssikkerheten. Med ukryptert database var det ikke behov for å utvikle og sette opp avanserte algoritmer for forespørsel og søk, men kryptering ved overføring av søkeresultatet var en mulighet for økt sikkerheten. Det antas derfor at denne modellen ville blitt den billigste løsningen å utvikle, og modellen får derfor en fordel på dette punktet. Men siden vurdering av kostnadsbegrensninger ble valgt bort til fordel for en sterkere informasjonssikkerhet rundt løsningen, falt valget på en løsning der man benyttet seg av et mer avansert tilgangssystem, og en kryptert database.

For en bedre oversikt har vi oppsummert de viktigste fordelene og ulempene med de ulike kryptoalgoritmene og den ukrypterte løsningen (Tabell 8.1).

I dette kapittelet har de sikkerhetsmessige og kostnadsmessige fordelene og ulempene ved valgt kryptoalgoritme for den presenterte lagringsløsningen blitt diskutert, og fordeler og ulemper ved alternative kryptoalgoritmer har blitt presentert for sammenligning. Utfordringer den presenterte modellen står ovenfor har blitt identifisert og vurdert, og eksterne og interne tiltak som kan redusere påvirkningen disse utfordringene har på modellens funksjonelle virkning har blitt presentert.

Tabell 8.1: Fordeler og ulemper med ulike kryptoalgoritmer.

Kryptoalgoritme	*Fordel*	*Ulempe*
Presentert modell med HVE	*Kryptert database.* *Samarbeid mellom mange entiteter for uthenting av data.*	*Størrelsesorden på PK, C_t og krypteringstid.*
IBE	*Kryptert database.* *Mindre komplisert enn HVE og MRQED.*	*Metode for tildeling av dekrypteringsrettigheter.* *Få innstanser nødvendig for vellykket innsideangrep.*
MRQED	*Kryptert database.* *Gunstig størrelsesorden på PK, C_t og krypteringstid.*	*Må vite alle parametre i tuppel.*
HVE med delegering	*Kryptert database.* *Samarbeid mellom mange entiteter for uthenting av data.*	*Unødvendige funksjoner, komplisert.*
Ukryptert database	*Enklere og billigere å utvikle.*	*Ukryptert database.*

Kapittel 9

Konklusjon

Dette arbeidet har presentert en modell for hvordan man kan lagre, sikre, og kontrollere tilgangen til, elektroniske kommunikasjonsdata som skal lagres fra 6 måneder til 2 år i følge EUs datalagringsdirektiv. Med bakgrunn i direktivteksten, norske lovtekster, akademiske artikler og publikasjoner, samt tilegnet kunnskap om dagens praksis for lagring og uthenting av elektroniske kommunikasjonsdata, har dette arbeidet skissert bruksmodeller og krav til hvordan en eventuell implementering av direktivet kan løses både på et administrativt og teknisk plan.

Mange hensyn har måttet tas da modellen for en lagringsløsning skulle utformes, inkludert kostnader, personvern og sikkerhet. Økte kostnader for tilbydere av elektronisk kommunikasjon har allerede vært nevnt som en konsekvens av en eventuell innføring av direktivet i Norge [3], [6]. Jo mer avansert den nye lagringsløsningen ville bli, jo lengre ble utviklingstiden, og jo dyrere ble en eventuell implementering. Av hensyn til et best mulig vern av brukernes privatliv ble sikkerheten rundt løsningen, både administrativt og teknisk, høyt prioritert. I dette arbeidet har sikkerheten blitt prioritert høyere enn de mulig tilhørende høye kostnadene både fordi dette arbeidets hovedfokus går på informasjonssikkerhet, og på grunn av løsningens effekt på personvernet.

De administrative kravene og den administrative løsningen inkluderte en implementering av grundige rutiner for personellsikkerhet, fysisk sikkerhet og prosedyrer rundt lagringsprosessen. Kravene ble basert på anbefalinger og standardiserte metoder fra ISO-standarder [16], [17], [18], [29] og en ETSI-standard [19]. De administrative prosessene inkluderte rutiner for håndtering av forholdene både utenfor systemet så vel som i selve systemet, for en best mulig forsikring om at dataene kun ble brukt til det de var ment for. Slike administrative rutiner eksisterer hos mange tilbydere av elektronisk kommunikasjon allerede i dag, da de allerede er underlagt taushetsplikt og norsk lovverk.

Etter å ha sett på flere alternative kryptoalgoritmer for bruk i den tekniske løsningen falt valget på Hidden Vector Encryption (HVE), en underalgoritme av Public key Encryption with Conjunctive Keyword Search (PECKS). Selv om denne kryptoalgoritmen hadde noe dårligere forutsetninger enn MRQED på kjøretid og størrelsesorden på parametere, ble HVE brukt i dette arbeidet på grunn av at man kunne bruke såkalte likegyldige verdier i søket etter data. Den presenterte tekniske modellen brukte HVE til å sikre at man skulle kunne kryptere de dataene som ble lagret, søke blant dataene uten å dekryptere innholdet, og til aksesskontroll. Kryptering og dekryptering skjedde med to forskjellige nøkler, PK og SK, og den skjulte nøkkelen, SK, ble brukt sammen et predikat, P, for å opprette spørringen mot databasen. Datafragmentene som trengtes for å søke etter, hente ut og lese data, ble fordelt mellom politiet, PT og retten. Dette for å kontrollere tilgangen til dataene samt redusere risikoen for at utro tjenere innenfor noen av disse instansene skulle kunne hente ut data. En mer sammensatt tilgangskontroll kan også redusere de effektene lagringen vil få på personvernet. Personvernet var også en av grunnene til at valget falt på å bruke en løsning med kryptert database fremfor en enklere, og muligens billigere, løsning der data ble lagret i klartekst.

Det ble ikke gitt noen garanti for at ulovlig anskaffelse av dataene i klartekst var umulig. Å tro at man skulle kunne garantere 100% sikkerhet gjennom

en solid teknisk løsning er både urealistisk og naivt. Sjansene for at trusselaktører kan gå rundt hindringene, eller bryte seg inn i den tekniske løsningen er absolutt til stede, men dette er også tilfelle for all annen lagring av hemmelige, offentlige eller private elektroniske data i dag. For å minimere sannsynligheten for at utro tjenere hos tilbyder, retten, PT og politiet skulle kunne hente ut data i klartekst, ble elementer som var nødvendige for aksess til databasen distribuert mellom nettopp disse entitetene.

Dette arbeidet har presentert en administrativ og teknisk modell for en ny lagringsløsning tilpasset en situasjon der datalagringsdirektivet eller lignende lov om lagring var innført i Norge. Informasjonssikkerhet har vært hovedfokus for modellen, og kostnader ved utvikling har vært et mindre viktig tema. Grundige administrative rutiner har derfor blitt utformet av flere ISO-standarder for håndtering av informasjonssikkerhet i organisasjoner, samt en ETSI-standard for håndtering av innsamling, lagring og bruk av elektroniske overvåkningsdata. I den tekniske delen har valget falt på å kryptere de lagrede dataene, og bruke en kryptoalgoritme kalt HVE til å håndtere kryptering og dekryptering, søk i krypterte data og aksesskontroll. HVE skjulte mest mulig av informasjonen i søket etter data, og tilgangskontrollen ble gjort sammensatt gjennom å distribuere kunnskap som trengtes for å aksessere systemet. Men modellen hadde ingen garanti for ikke at innsideangrep ikke var mulig.

Bibliografi

[1] Europa-Parlamentet og Rådet for Den Europeiske Union. Direktiv 2006/24/EF. [Online], <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2006:105:0054:0063:DA:PDF>, 2006.

[2] Regjeringen. Høringsnotat - Datalagring. [Online], <http://www.regjeringen.no/dep/sd/dok/hoeringer/hoeringsdok/2010/horing—datalagring/horingsnotat.html?id=590003>, 2010.

[3] Teleplan v/Kristin Kaarstad. Økonomisk utredning av konsekvensene knyttet til innføring av EUs datalagringsdirektiv. Technical report, Teleplan, 2008.

[4] Justis-og politidepartementet. Lov-2000-04-14-31 lov om behandling av personopplysninger (personopplysningsloven). Publisert i 2000 hefte 8, 2000.

[5] Samferdselsdepartementet. Lov-2003-07-04-83 lov om elektronisk kommunikasjon (ekomloven). Publisert i 2003 hefte 10, 2003.

[6] Teleplan v/Kristin Kaarstad. Økonomisk utredning av konsekvensene knyttet til innføring av EUs datalagringsdirektiv. Technical report, Teleplan, 2006.

[7] Justis-og politidepartementet. Lov-1981-05-22-25 lov om rettergangsmåten i straffesaker (straffeprosessloven). ISBN 82-504-1306-7, 1986.

[8] Svein Willassen. Datalagringsdirektivet - Verdi i Etterforskning og Risikofaktorer for Personvern. *for Datatilsynet*, 2010.

[9] Svein Willassen. Å lagre eller ikke å lagre. *Lov&Data*, (94), 2008.

[10] Arne Johannesen og Anne-Catherine Gustafson. NOTAT - EU's datalagringsdirektiv. [Online], <http://www.pf.no/id/15774>, 2009.

[11] Stopp Datalagringsdirektivet. Høring om datalagringsdirektivet. [Online] <http://stoppdld.no/2010/04/13/her-erh%C3%B8ringsuttalelsen-fra-stopp-dld/>, 2010.

[12] Post-og teletilsynet. Registrerte tilbyder av offentlig ekomnett og offentlig teletjenester. Technical report, Post- og teletilsynet, 2010.

[13] Forsvarsdepartementet. Lov-1998-03-20-10 lov om forebyggende sikkerhetstjeneste (Sikkerhetsloven). Avd I 1998 Nr. 5, 2001.

[14] Europa-Parlamentet og Rådet for Den Europeiske Union. Direktiv 95/46/EF. [Online], <http://ec.europa.eu/justice_home/fsj/privacy/docs/95-46-ce/dir1995-46_part1_da.pdf>, 1995.

[15] Fornyings-Administrasjons-og Kirkedepartementet Artikkel 29 gruppen. NOU 2009: 1 - Uttalelse om Datalagringsdirektivet. [Online], <http://www.regjeringen.no/nb/dep/fad/dok/nouer/2009/nou-2009-1/24.html?id=542291#>, 2008.

[16] ISO/IEC. ISO/IEC 15408-1: Evaluation Criteria for IT Security – Part 1: Introduction and general model. Technical report, Common Criteria, 2009.

[17] ISO/IEC. ISO/IEC 27000, Information Technology - Security Techniques - Information Security Management Systems. Technical report, International Organization of Standardization & International Electrotechnical Commission, 2009.

[18] ISO/IEC. ISO/IEC 27002:2005, Information Technology - Security Techniques - Code of Practice for Information Security Management. Technical report, International Organization of Standardization & International Electrotechnical Commission, 2009.

[19] ETSI TR 102 661. Lawful Intercetion (LI); Security framework in Lawful Interception and Retained Data environment. Technical report, ETSI, 2008.

[20] Adi Shamir. Identity-based Cryptosystems and Signature Schemes. *Lecture Notes in Computer Science*, pages 47–53, 1985.

[21] Dong Jin Park, Kihyun Kim, and Pil Joong Lee. Public key encyprtion with conjunctive field keyword search. *Lecture Notes in Computer Science*, pages 73–86, 2004.

[22] Vincenzo Ivoino and Giuseppe Persiano. Hidden-Vector Encryption with Groups of Prime Order. *Lecture Notes in Computer Science*, pages 75–88, 2007.

[23] Elaine Shi and Brent Waters. Delegating Capabilities in Predicate Encryption Systems. *Lecture Notes in Computer Science*, pages 560–578, 2008.

[24] Elaine Shi, John Bethencourt, T-H. Hubert Chan, Dawn Song, and Adrian Perrig. Multi-Dimentional Range Query over Encrypted Data. *Carnegie Mellon University*, 2007.

[25] Dan Boneh and Brent Waters. Conjunctive, Subset and Range Queries on Encrypted Data. *Lecture Notes in Computer Science*, pages 535–554, 2007.

[26] Dan Boneh and Matthew Franklin. Identity-Based Encryption from the Weil Pairing. *Lecture Notes in Computer Science*, 2139:213–229, 2003.

[27] Clifford Cooks. An Identity Based Encryption Scheme based on Quadratic Residues. *Communications-Electronics Security Group*, 2001.

[28] Dag-Rune Z. Vollen. Lagringsdirektivet kan koste milliarder. [Online], <http://www.idg.no/computerworld/article167195.ece?curPage=> [aksessert 18.05.10], Computerworld, 2010.

[29] ISO/IEC. ISO/IEC 15408-2: Evaluation Criteria for IT Security – Part 2: Security functional requirements. Technical report, Common Criteria, 2009.

Vedlegg

Den sammensatte 3-delte Diffie-Hellman antakelsen (C3DH)

Sikkerheten i HVE bruker den sammensatte 3-delte Diffie-Hellman antakelsen (composite 3-party Diffie-Hellman assumption) (C3DH).

For en gitt gruppe $\mathcal{G}$ defineres følgende fordeling $P(\lambda)$:

$(p,\, q,\, \mathbb{G},\, \mathbb{G}_T,\, e) \leftarrow^R \mathcal{G}(\lambda), \quad n \leftarrow pq, \quad g_p \leftarrow^R \mathbb{G}_p, \quad g_q \leftarrow^R \mathbb{G}_q$

$R_1, R_2, R_3, \leftarrow^R \mathbb{G}_q$

$a,\, b,\, c \leftarrow^R \mathbb{Z}_n$

$Z \leftarrow ((n, \mathbb{G}, \mathbb{G}_T, e), g_q, g_p, g_p^a, g_p^b, g_p^{ab} \cdot R_1, g_p^{abc} \cdot R_2)$

$T \leftarrow g_p^c \cdot R_3$

Output $(Z,\, T)$

For en algoritme $\mathcal{A}$, defineres $\mathcal{A}$s fordel ved å løse den samensatte 3-delte Diffie-Hellman antakelsen for problem $\mathcal{G}$ som:

$\text{C3DH}Adv_{\mathcal{G},\mathcal{A}}(\lambda) := |Pr[\mathcal{A}(Z,T) = 1] - Pr[\mathcal{A}(Z,R) = 1]|$

der $(Z,\ T) \leftarrow^R P(\lambda)$ og $R \leftarrow^R \mathbb{G}$

Vi sier da at $\mathcal{G}$ tilfredsstiller den sammensatte 3-delte Diffie-Hellman antakelsen (C3DH) hvis for hvilken som helst algoritme $\mathcal{A}$ av polynomisk tidsorden vi får at $\text{C3DH}Adv_{\mathcal{G},\mathcal{A}}(\lambda)$ er en ubetydelig funksjon av λ.