Koalisjons-sikker fingerprenting.
Ei simulering på Boneh-Shaw systemet.

Koalisjons-sikker fingerprenting.
Ei simulering på Boneh-Shaw systemet.

Forelesninga:

1. Introduksjon til fingerprenting.
2. Boneh-Shaw systemet.
3. Mine resultat.
4. Framtidig arbeid.

Kvifor forska på fingerprentingsfeltet?

- Ulovleg distribusjon er eit stort problem.
  Digitale produkt spesielt utsett fordi:
  - Enkel kopiering.
  - Inga informasjon går tapt.
  - Internett mogleggjer storskaladistribusjon.
  - Ofte vanskeleg å finna den ansvarlege.
  - Ofte vanskeleg å få dømd den ansvarlege.

Kvifor forska på fingerprentingsfeltet?

- Ulovleg distribusjon er eit stort problem.
  Digitale produkt spesielt utsett fordi:
  - Enkel kopiering.
  - Inga informasjon går tapt.
  - Internett mogleggjer storskaladistribusjon.
  - Ofte vanskeleg å finna den ansvarlege.
  - Ofte vanskeleg å få dømd den ansvarlege.
- Kopisperren på digitale produkt blir som regel broten.

Så kva er fingerprenting?

- Deteksjonsteknikk.
  - Mogleggjer sporing av ein kjøpar som har distribuert ulovleg.
  - Målet er å skremma kjøparar mot å spreie eit produkt vidare.

Så kva er fingerprenting?

- Deteksjonsteknikk.
  - Mogleggjer sporing av ein kjøpar som har distribuert ulovleg.
  - Målet er å skremma kjøparar mot å spreie eit produkt vidare.

Fingerprenting i praksis:

- Unikt merke i kvar kopi av produktet.
- Knyttar kvar kjøpar opp mot ein unik kopi.
- Om kopien vert funnen veit distributør kven som er skuldig (piraten).

Så kva er fingerprenting?

- Deteksjonsteknikk.
  - Mogleggjer sporing av ein kjøpar som har distribuert ulovleg.
  - Målet er å skremma kjøparar mot å spreie eit produkt vidare.

Fingerprenting i praksis:

- Unikt merke i kvar kopi av produktet.
- Knyttar kvar kjøpar opp mot ein unik kopi.
- Om kopien vert funnen veit distributør kven som er skuldig (piraten).

Digital fingerprenting introdusert i 1983 (Wagner).

Så kva er fingerprenting?

- Deteksjonsteknikk.
  - Mogleggjer sporing av ein kjøpar som har distribuert ulovleg.
  - Målet er å skremma kjøparar mot å spreie eit produkt vidare.

Fingerprenting i praksis:

- Unikt merke i kvar kopi av produktet.
- Knyttar kvar kjøpar opp mot ein unik kopi.
- Om kopien vert funnen veit distributør kven som er skuldig (piraten).

Digital fingerprenting introdusert i 1983 (Wagner).

Døme 1 *Logaritmetabellar.*

- *Feil blir introdusert.*
- *Viktig at desse feila er unike for kvar kopi.*

Logaritmetabell. (Døme: $\log 109,7 = 2,040207$))

	0	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
100	.000000	.000434	.000868	.001301	.001734	.002166	.002598	.003029	.003461	.003891
101	.004321	.004751	.005181	.005609	.006038	.006466	.006894	.007321	.007748	.008174
102	.008600	.009026	.009451	.009876	.010300	.010724	.011147	.011570	.011993	.012415
103	.012837	.013259	.013680	.014100	.014521	.014941	.015360	.015779	.016197	.016616
104	.017033	.017451	.017868	.018284	.018701	.019116	.019532	.019947	.020361	.020776
105	.021189	.021603	.022016	.022428	.022841	.023253	.023664	.024075	.024486	.024896
106	.025306	.025715	.026125	.026533	.026942	.027350	.027757	.028164	.028571	.028978
107	.029384	.029790	.030195	.030600	.031004	.031409	.031813	.032216	.032619	.033022
108	.033424	.033826	.034227	.034628	.035029	.035430	.035830	.036230	.036629	.037028
109	.037427	.037825	.038223	.038620	.039017	.039414	.039810	.040207	.040603	.040998
110	.041393	.041787	.042182	.042576	.042969	.043363	.043775	.044148	.044540	.044932
...										

Logaritmetabell. (Døme: $\log 109,7 = 2,040207$))

	0	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
100	.000000	.000434	.000868	.001301	.001734	.002166	.002598	.003029	.003461	.003891
101	.004321	.004751	.005181	.005609	.006038	.006466	.006894	.007321	.007748	.008174
102	.008600	.009026	.009451	.009876	.010300	.010724	.011147	.011570	.011993	.012415
103	.012837	.013259	.013680	.014100	.014521	.014941	.015360	.015779	.016197	.016616
104	.017033	.017451	.017868	.018284	.018701	.019116	.019532	.019947	.020361	.020776
105	.021189	.021603	.022016	.022428	.022841	.023253	.023664	.024075	.024486	.024896
106	.025306	.025715	.026125	.026533	.026942	.027350	.027757	.028164	.028571	.028978
107	.029384	.029790	.030195	.030600	.031004	.031409	.031813	.032216	.032619	.033022
108	.033424	.033826	.034227	.034628	.035029	.035430	.035830	.036230	.036629	.037028
109	.037427	.037825	.038223	.038620	.039017	.039414	.039810	.040207	.040603	.040998
110	.041393	.041787	.042182	.042576	.042969	.043363	.043775	.044148	.044540	.044932
...										

Modifisert logaritmetabell.

	0	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
100	.000000	.000434	.000868	.001301	.001734	.002166	.002598	.003029	.003461	.003891
101	.004321	.004751	.005181	.005609	.00603**9**	.006466	.006894	.007321	.007748	.008174
102	.008600	.009026	.009451	.009876	.010300	.010724	.011147	.011570	.011993	.012415
103	.012837	.013259	.013680	.014100	.014521	.014941	.015360	.015779	.016197	.016616
104	.017033	.017451	.017868	.018284	.01870**2**	.019116	.019532	.019947	.020361	.020776
105	.021189	.021603	.022016	.022428	.022841	.023253	.023664	.024075	.024486	.024896
106	.025306	.025715	.02612**4**	.026533	.026942	.027350	.027757	.028164	.028571	.028978
107	.029384	.029790	.030195	.030600	.031004	.031409	.031813	.032216	.032619	.033022
108	.033424	.033826	.034227	.034628	.035029	.035430	.035830	.036230	.036629	.037028
109	.037427	.037825	.038223	.038620	.039017	.039414	.039810	.040207	.040603	.040998
110	.041393	.041787	.042182	.042576	.042969	.043363	.043775	.044148	.044540	.044932
...										

Koalisjons-sikker fingerprenting.

Problem førekjem når piratane samarbeidar om å laga eit hybridfingeravtrykk.

	0	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
100	.000000	.000434	.000868	.001301	.001734	.002166	.002598	.003029	.003461	.003891
101	.004321	.004751	.005181	.005609	.00603**9**	.006466	.006894	.007321	.007748	.008174
...										

	0	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
100	.000000	.000434	.000868	.001301	.001734	.002166	.002598	.003029	.003461	.003891
101	.004321	.004751	.005181	.005609	.00603**8**	.006466	.006894	.007321	.007748	.008174
...										

Koalisjons-sikker fingerprenting.

Problem førekjem når piratane samarbeidar om å laga eit hybridfingeravtrykk.

	0	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
100	.000000	.000434	.000868	.001301	.001734	.002166	.002598	.003029	.003461	.003891
101	.004321	.004751	.005181	.005609	.00603**9**	.006466	.006894	.007321	.007748	.008174
...										

	0	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
100	.000000	.000434	.000868	.001301	.001734	.002166	.002598	.003029	.003461	.003891
101	.004321	.004751	.005181	.005609	.00603**8**	.006466	.006894	.007321	.007748	.008174
...										

Koalisjons-sikker fingerprenting blir brukt for å motverka dette problemet.

Relaterte felt.

- Vassmerking.
  - Kan bevisa eigarskap.
  - Alle kopiar er identisk vassmerka.
  - Kan ikkje spora skyldig brukar.
  - Bør vere robust mot angrep.
- Fingerprenting.
  - Målet er å forhindra spreiing av kopibeskytta matriell.
  - Alle selde objekt har eit unikt fingeravtrykk.
  - Koalisjon lagar nytt fingeravtrykk.
  - Mogleg å spore koalisjon med høg sannsynlegheit.
  - Treng ein kopi av objektet for sporing.
- Forrædarsporing (Traitor tracing).
  - Målet er å forhindra spreiing av dekrypteringsnøklar.
  - Kvar brukar får ein unik dekrypteringsnøkkel.
  - Koalisjon lagar ny nøkkel.
  - Mogleg å spore koalisjon med høg sannsynlegheit.
  - Treng kopi av dekrypteringsnøkkel for sporing.

Dei fem delane i fingerprenting.

1. Kodar. Eit fingeravtrykk er eit kodeord.
2. Embedding. Fingeravtrykket lyt setjast inn i objektet.
3. Lenking. Fingeravtrykk lyt knyttast til kjøpar.
4. Piratstrategiar. Strategiar for generering av hybridfingeravtrykk.
5. Sporing. Identifiserar piratane.

Kodar

$C = (n, M)_q$-kode.

- n er lengda på fingeravtrykka.
- q er storleiken på kodealfabetet.
- M er talet på kjøparar støtta av koden.

Døme 2

Kodar

$C = (n, M)_q$-kode.

- n er lengda på fingeravtrykka.
- q er storleiken på kodealfabetet.
- M er talet på kjøparar støtta av koden.

Døme 2

La C vera ein kode med 4 kodeord, lengd 5 og alfabetstørrelse 2. C er då ein $(5, 4)_2$-kode.

$$C = \begin{cases} 11000 \\ 01100 \\ 00110 \\ 00011 \end{cases}$$

Kodar

$C = (n, M)_q$-kode.

- n er lengda på fingeravtrykka.
- q er storleiken på kodealfabetet.
- M er talet på kjøparar støtta av koden.

Døme 2

La C vera ein kode med 4 kodeord, lengd 5 og alfabetstørrelse 2. C er då ein $(5, 4)_2$-kode.

$$C = \begin{cases} 11000 \\ 01100 \\ 00110 \\ 00011 \end{cases}$$

Kodeorda i C formar ei matrise som me kallar kodeboka. Frå denne kodeboka gjev ein rad i til kjøpar i, der $1 \leq i \leq M$.

Konkatenering av kodar.

Komponentar:

- Innerkode $C_1 = (n_1, q)$.
- Ytterkode $C_2 = (n_2, M)_q$.

Konkatenering av kodar.

Komponentar:

- Innerkode $C_1 = (n_1, q)$.
- Ytterkode $C_2 = (n_2, M)_q$.

Resulterar i den konkatenerte koden:

- $C' = (n_1 n_2, M)_q$.

  - q finn ein i begge kodane.
  - Den konkatenerte koden er samansetjinga av innerkodeorda som ytterkodesymbola peikar til.

Døme 3 *Konkatenering.*
$C_2 = (n_2, M)_q = (4,4)_4$ *ytterkode*

$$C_2 = \begin{cases} 1234 \\ 2341 \\ 3412 \\ 4123 \end{cases}$$

$C_1 = (n_1, q) = (4,4)$ *innerkode*

$$C_1 = \begin{cases} 1000 \\ 0100 \\ 0010 \\ 0001 \end{cases}$$

Døme 3 *Konkatenering.*
$C_2 = (n_2, M)_q = (4,4)_4$ *ytterkode*

$$C_2 = \begin{cases} 1234 \\ 2341 \\ 3412 \\ 4123 \end{cases}$$

$C_1 = (n_1, q) = (4,4)$ *innerkode*

$$C_1 = \begin{cases} 1000 \\ 0100 \\ 0010 \\ 0001 \end{cases}$$

$C' = (n_1 n_2, M)_q = (16,4)_4$ *konkatenert kode*

$$C' = \begin{cases} 1000||0100||0010||0001 \\ 0100||0010||0001||1000 \\ 0010||0001||1000||0100 \\ 0001||1000||0100||0010 \end{cases}$$

Den oppnåelege mengda.

Har ein piratkoalisjon P av størrelse t.

For P der $t \geq 2$ er den oppnåelege mengda dei kodeord P kan generera.

Døme 4 *Lat A og B vere to brukarar som har kjøpt objekt med kodeorda:*

Brukar A	=	10110
Brukar B	=	10011

Den oppnåelege mengda.

Har ein piratkoalisjon P av størrelse t.

For P der $t \geq 2$ er den oppnåelege mengda dei kodeord P kan generera.

Døme 4 *Lat A og B vere to brukarar som har kjøpt objekt med kodeorda:*

Brukar A	=	10110
Brukar B	=	10011
Den oppnåelege mengda	=	10?1?

Den oppnåelege mengda.

Har ein piratkoalisjon P av størrelse t.

For P der $t \geq 2$ er den oppnåelege mengda dei kodeord P kan generera.

Døme 4 *Lat A og B vere to brukarar som har kjøpt objekt med kodeorda:*

Brukar A	=	10110
Brukar B	=	10011
Den oppnåelege mengda	=	10?1?
Hybridfingeravtrykk	=	10010

Den oppnåelege mengda.

Har ein piratkoalisjon P av størrelse t.

For P der $t \geq 2$ er den oppnåelege mengda dei kodeord P kan generera.

Døme 4 *Lat A og B vere to brukarar som har kjøpt objekt med kodeorda:*

Brukar A	=	10110
Brukar B	=	10011
Den oppnåelege mengda	=	10?1?
Hybridfingeravtrykk	=	10010

Merkningsføresetnaden.

Definerar kva ein pirat har lov til å gjera.

Piratstrategiar.

Strategiar for å setja detekterte symbol.

Dette kan vere ukjent for piratane:

1. Udetekterte symbol i objektet (alltid).
2. Korrespondansen mellom symbola i objektet og symbola i kodeordet.
3. Kva for kodeord som er gitt til dei forskjellige brukarane.

Piratstrategiar.

Strategiar for å setja detekterte symbol.

Dette kan vere ukjent for piratane:

1. Udetekterte symbol i objektet (alltid).
2. Korrespondansen mellom symbola i objektet og symbola i kodeordet.
3. Kva for kodeord som er gitt til dei forskjellige brukarane.

Deterministiske og tilfeldige strategiar:

1. Deterministisk strategiar:
    - Majoritetsvalg. Det symbolet som førekjem flest gonger i objekta til piratkoalisjonen blir valgt.
    - Minoritetsvalg: Motsatt av majoritetsvalg.
2. Tilfeldig strategi:
    - Velja tilfeldig blant alle dei symbola ein kan sjå, med lik sannsynlegheit.

Sporing av piratar.

Ei sporingsalgoritme tek eit hybridfingeravtrykk som inndata og gjev ei delmengd $P \subseteq C$ som utdata.

Probabilistisk og kombinatorisk:

- Kombinatoriske skjema gjev alltid ei delmengd av P som utdata.
- Probabilistiske skjema gjev ei delmengd av P med sannsynlegheit minst $1 - \epsilon$, der ϵ er ein låg feilrate.

Sporing av piratar.

Ei sporingsalgoritme tek eit hybridfingeravtrykk som inndata og gjev ei delmengd $P \subseteq C$ som utdata.

Probabilistisk og kombinatorisk:

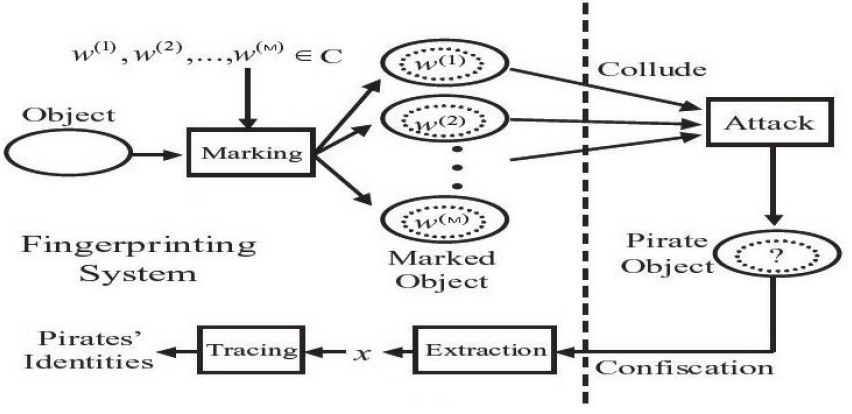
- Kombinatoriske skjema gjev alltid ei delmengd av P som utdata.
- Probabilistiske skjema gjev ei delmengd av P med sannsynlegheit minst $1 - \epsilon$, der ϵ er ein låg feilrate.

Sporingsalgoritma i eit probabilistisk skjema kan gi tre resultat:

1. Vellukka sporing: Utdata er ei ikkje-tom delmengd av P.
2. Type I feil: Utdata er ei tom mengd.
3. Type II feil: Utdata inneheld ein eller fleire uskuldige kjøparar.

Type II er mest alvorleg.

Skissering av eit fingerprentingssystem



Boneh-Shaw systemet

Kjend koalisjons-sikkert system.

- Konkatenert kode.
- Sporingsalgoritmen returnerar alltid 1 brukar.
- Probabilistisk.
  - Gitt feilsannsynlegheit ϵ, antall brukarar M og antallet piratar t kan ein velgja kodeordlengda slik at den verkelege feilsannsynlegheita er mindre enn ϵ.

t-sikre kodar med ϵ-feil.

Definisjon 1 *Eit fingerprentingssystem er t-sikkert med ϵ-feil om ein kan spora ein piratkoalisjon, med høgst t piratar, med sannsynlegheit bedre enn $1 - \epsilon$.*

- Probabilistisk skjema.
- Permutasjon på kodeorda før embeddinga.

Replikasjonskoden.

Har ein kode: $\Gamma(q, r) = \Gamma(4, 3)$. $(n_1 = r(q-1))$.

$$\Gamma = \begin{cases} 111111111 \\ 000111111 \\ 000000111 \\ 000000000 \end{cases}$$

Kvar kolonnetype replikert r gonger.

Replikasjonskoden.

Har ein kode: $\Gamma(q, r) = \Gamma(4, 3)$. $(n_1 = r(q-1))$.

$$\Gamma = \begin{cases} 111111111 \\ 000111111 \\ 000000111 \\ 000000000 \end{cases}$$

Kvar kolonnetype replikert r gonger.

- Permutasjon på kodeorda.
- For ein koalisjon uten kodeord 2 ser dei 6 første posisjonane like ut.
- Fordela jamt.
- Ikkje jamt, då er kodeord 2 med i P med stor sannsynlegheit.
- Sporingsalgoritma bruker denne informasjonen.

Permutert kode:

$$\Gamma_p = \begin{cases} \underline{111111}111 \\ 101010111 \\ \underline{000000}111 \\ \underline{000000}000 \end{cases}$$

Det konkatenerte skjemaet

Probabilistisk t-sikker kode med logaritmisk lengd.

Komponentar:

1. Innerkode: Replikasjon. $\Gamma(q, r)$.
2. Ytterkode: Tilfeldig. $C(n_2, M)_q$.

Den konkatenerte koden: $C'(n_1 n_2, M)_q$, der $n_1 = r(q-1)$.

- Gitt M, t og $\epsilon > 0$, kan ein kalkulera størrelsen på q, r og n_2.
- Den konkatenerte koden er då t-sikker med ϵ-feil.

Det konkatenerte skjemaet

Probabilistisk t-sikker kode med logaritmisk lengd.

Komponentar:

1. Innerkode: Replikasjon. $\Gamma(q, r)$.
2. Ytterkode: Tilfeldig. $C(n_2, M)_q$.

Den konkatenerte koden: $C'(n_1 n_2, M)_q$, der $n_1 = r(q-1)$.

- Gitt M, t og $\epsilon > 0$, kan ein kalkulera størrelsen på q, r og n_2.
- Den konkatenerte koden er då t-sikker med ϵ-feil.

Sporingsalgoritma:

1. Innerkodealgoritma køyrt for kvart innerkodeord.
2. Set saman ytterkodesymbola.
3. Næraste granne-dekoding (closest neighbour decoding).
4. Returner brukar med flest felles posisjonar som skyldig.

Bakgrunn for oppgåva mi.

- Kun teoretiske grenser var kjent.
- Samanlikne teoretiske grenser med empiriske resultat.
- Køyra simuleringar på Boneh-Shaw systemet.
- Teoretiske grenser upresise.
- Forandrar størrelse på kodeparametrar og studerar feilratane og køyretidene.

Bakgrunn for oppgåva mi.

- Kun teoretiske grenser var kjent.
- Samanlikne teoretiske grenser med empiriske resultat.
- Køyra simuleringar på Boneh-Shaw systemet.
- Teoretiske grenser upresise.
- Forandrar størrelse på kodeparametrar og studerar feilratane og køyretidene.

Korleis måle ytelsen i eit fingerprentingssystem.

- Kor høg er sannsynlegheita for å spore piratane?
- Kor korte fingeravtrykk kan ein bruke?
- Kor stor piratkoalisjon taklar ein?
- Kor gode køyretider førekjem?

Bakgrunn for oppgåva mi.

- Kun teoretiske grenser var kjent.
- Samanlikne teoretiske grenser med empiriske resultat.
- Køyra simuleringar på Boneh-Shaw systemet.
- Teoretiske grenser upresise.
- Forandrar størrelse på kodeparametrar og studerar feilratane og køyretidene.

Korleis måle ytelsen i eit fingerprentingssystem.

- Kor høg er sannsynlegheita for å spore piratane?
- Kor korte fingeravtrykk kan ein bruke?
- Kor stor piratkoalisjon taklar ein?
- Kor gode køyretider førekjem?

Piratstrategien brukt i simuleringa er ei tilfeldig, sjølvstendig generering av binære tal for alle detekterte posisjonar.

Kodeparametrar.

Viser korleis ei endring i q, n_2 eller r påverkar ytelsen.

Kodeparameter	n	ϵ	Sporingstid
Auke i q	$\uparrow$	$\downarrow$	$\uparrow$
Auke i n_2	$\uparrow$	$\downarrow$	$\uparrow$
Auke i r	$\uparrow$	$\searrow$	$\uparrow$

Viktigaste resultat.

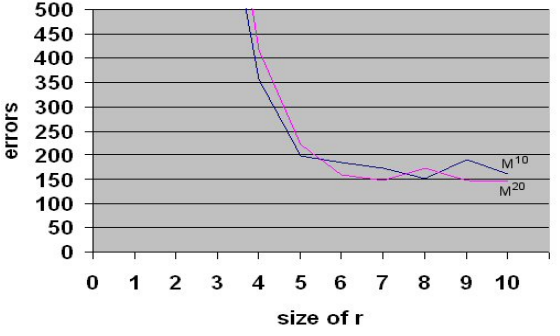
Replikasjonsfaktoren r er unødvendig stor.

- For Boneh-Shaw aukar r ettersom ϵ blir bedre.
- Døme: $\epsilon = 10^{-10}$, $t = 10$ og $M = 2^t$ gjev $r = 39465$.

Viktigaste resultat.

Replikasjonsfaktoren r er unødvendig stor.

- For Boneh-Shaw aukar r ettersom ϵ blir bedre.
- Døme: $\epsilon = 10^{-10}$, $t = 10$ og $M = 2^t$ gjev $r = 39465$.
- Vist at r bør vere ca. 5.



- Gjev kortare kode og kortare køyretid.

Fleire resultat.

Held n konstant, forandrar n_2 og q.

n_2	q	Feil	Feilrate	Køyretid
1000	6	79680	$1.59 \cdot 10^{-2}$	13.37 timar
500	11	157	$3.14 \cdot 10^{-5}$	6.31 timar
250	21	10	$2 \cdot 10^{-6}$	3.35 timar
125	41	12	$2.4 \cdot 10^{-6}$	2.16 timar
40	126	864	$2.73 \cdot 10^{-4}$	1.18 timar
20	251	14263	$2.85 \cdot 10^{-3}$	1.06 timar
10	501	113384	$2.26 \cdot 10^{-2}$	0.59 timar
5	1001	829917	$1.65 \cdot 10^{-1}$	0.55 timar

Her er $t = 10$, $M = 2^t$, $r = 5$, totalt 5000000 køyringar.

n_2	q	Feil	Feilrate	Køyretid
1000	21	5171	$1.03 \cdot 10^{-1}$	113.23 timar
500	41	640	$1.28 \cdot 10^{-2}$	59.66 timar
250	81	959	$1.91 \cdot 10^{-2}$	26.39 timar
160	126	2027	$4.05 \cdot 10^{-2}$	17.37 timar
125	161	3550	$7.1 \cdot 10^{-2}$	14.23 timar
80	251	7674	$1.53 \cdot 10^{-1}$	9.42 timar
40	501	18912	$3.78 \cdot 10^{-1}$	5.58 timar
20	1001	29100	$5.82 \cdot 10^{-1}$	4.11 timar

Her er $t = 20$, $M = 2^t$, $r = 2$, totalt 50000 køyringar.

Opne problem.

- Samanlikne mine resultat mot teoretiske grenser.
- Samanlikne ulike skjema.
- Samanlikne fleire piratstrategiar.