Mail reloaded: Censorship-resistant high-latency mail communication

Joakim Grebenö

1 september 2020

Abstract

There is a theory which states that if ever anyone discovers exactly what the Universe is for and why it is here, it will instantly disappear and be replaced by something even more bizarre and inexplicable. There is another theory mentioned, which states that this has already happened.[1]

1 Bakgrund

Anonymitet och tekniker mot censur är två aktiva forskningsområden, och det blir extra tydligt i forskningsöversikter [15][16] och i de artiklar de refererar till. Det har hänt mycket sedan 2012 då jag arbetade med en implementation [3] av Schlegel och Wongs anonyma routingprotokoll [11][12] [13].

Min implementation står fortfarande ofärdig men Schlegel och Wong beskrev något som jag tycker fortfarande har potential. Så här i efterhand kan jag konstatera att de i praktiska termer var för ofullständiga och att de försökte lösa ett för generellt problem. De lämnade många praktiska problem olösta, t ex behovet av en Public Key Infrastructure (PKI) samt kravet på att kompletta routingtabeller måste synkroniseras mellan alla noder i nätverket. Dessutom fokuserade de på låglatent anonymitet som per definition kräver att det görs avkall på graden av anonymitet [22].

Forskning runt anonymitet och tekniker mot censur känns ännu mera som en bikupa än det gjorde 2012 och det är inte orimligt med tanke på den massövervakning som sedan dess implementerats av t ex Great Firewall of China (GFW) [1] och National Security Agency (NSA) [2].

1 Douglas Adams, **The Restaurant at the End of the Universe**, from the radio series, on Christmas Eve, 1978

Det är inte utan att jag ödmjukas av den mängd forskning som presenteras i Free Haven's Selected Papers in Anonymity [4] och i Selected Research Papers in Internet Censorship [5] men även i det outtröttliga arbete som pågår inom ramen för Tor-projektet [6].

1.1 Chaum-Mixes

David L. Chaum publicerade 1981 ”Untraceble Electronic, Return Addresses, and Digital Pseudonyms" [54] vars teorier om mixnätverk legat till grund för de flesta efterkommande anonymitetsnätverk. Tor [14] är ett exempel på ett mixnätverk som använder lökrouting [55].

Mixnätverk baseras på ett routingprotokoll som använder en kedja av proxyservrar/mixes, som tar emot och mixar data från många avsändare; data som sedan skickas vidare till nästa proxyserver i kedjan etc.

Mixnätverk gör det svårt för utomstående att kontrollera vem som kommunicerar med vem, och för att lyckas med detta användes assymetrisk kryptering som introducerades av kryptologen Whitfield Diffie i mitten på 70-talet.

1.2 Anonyma remailers

Anonyma remailers [31] introducerades 1992 och var en av de första allmänt spridda tillämpningar som baserades på mixnätverk. Först ut var Cypherpunk (typ I) remailern [32] som på grund av en svag anonymitet ersattes av Mixmaster (typ II) [33], som i sin tur ersattes av Mixminion (typ III) [34].

Värt att notera är att personerna bakom Mixminion är Roger Dingledine och Nick Mathewson som nu är centrala i utvecklingen av Tor.

Anonyma remailers var populära under nästan femton års tid men tappade relevans trots att behovet av stark anonymitet inte hade minskat; speciellt inte hos dissidenter. Anledningen till remailerdöden var troligen multivariabel:

a. Förhoppningarna på Tor var stora
b. Massövervakning var inte en realitet och mail användes inte av alla
c. Medvetenheten om säkerhetsbristerna hos mail var inte allmänt kända
d. Användningen av remailers krävde en viss teknisk kompetens och insikt
e. Remailers användes främst för att skicka anonyma mail till icke anonyma mottagare och detta ledde till uppmärksammade förekomster av stalking/spam (om så begränsad i mängd)

f. Dubbelriktad anonym mail-konversation kunde inte massövervakas av säkerhetstjänster och priset var högt för de som driftade remailers/nym-servrar [35]

Många av dessa anledningar är inte längre gällande: Tor har inte uppfyllt förhoppningarna (a) på stark anonymitet, främst på grund av kravet på låg latens [22]. (b) och (c) är inte längre sanna och (e) är inte relevant om man enbart fokuserar på (f). (d) stämmer inte om man anammar mail-metaforen och låter användare använda sina mail-läsare as-is.

Historiskt må anonyma remailers ha erbjudit någon grad av anonymitet men de var lätta att blockera på IP/DNS-nivå då de förlitade sig på centrala nym-servrar. De noder som ingick i dessa mixnätverk var också relativt lätta att blockera.

1.3 Tor: Det största mixnätverket

Tor startade som ett projekt finansierat av Defense Advanced Research Projects Agency (DARPA) och är fortfarande beroende av finansiering från en mängd organisationer [55]. Projektets uthållighet över tid har dock resulterat i ett mixnätverk som betjänar ~2 miljoner användare. Tor består av ~8000 frivillignoder som till ~98% (400GB/s) används för anonym surfning till publika siter [20]. Resterande användning handlar om access till Onion Services [18][21] på Darknet [17].

Värt att notera är att Tors protokollspecifikation är så pass föränderlig och ouppdaterad att det bara finns en komplett implementation och det är den egna referensimplementationen.

Det finns alternativa mixnätverk, såsom Freenet [27] och I2P [28], men de lever en marginaliserad tillvaro.

Ett flertal företag har försökt bygga kommersiella mixnätverk, t ex Anonymizer [29] och Silent Circle [30], men ingen av dem har har lyckats och de lever i bästa fall en tynande tillvaro. I maj 2019 drog dock Nym Technologies [7] in $2.5 million i riskkapital [8] för att bygga ett nytt kommersiellt mixnätverk. Nyms utvecklingsteam har bakgrund från implementationen av Loopix [9] och det förklarar deras val av tekniker:

- Probabilistic possion mixing with timing obfuscation
- Sphinx packet format
- Cover traffic
- Blockchains for identity management via Coconut-signaturesSystem rewards to incentivize node owners to perform CPU intensive mixing, i.e. to perform mixing instead of solving Merke puzzles during Bitcoin mining

Gemensamt för både Tor och Nym (och de flesta andra aktiva anonymitetsprojekt) är fokuset på låglatent anonymitet och priset har hitills alltid varit en svagare grad av anonymitet [22]. Låg latens är rimligen ett krav för de som vill surfa anonymt men dissidenter, som har ett större behov av just stark anonymitet, torde ha högre tolerans mot hög latens.

Ett problem för Tor är att klienter måste ladda ner en lista över alla noder som ingår i löknätverket (IP-adresser och publika nycklar etc) innan de kan börja kommunicera anonymt. Dessa listor tillhandahålls av ett tiotal statiska biblioteksservrar och utgör en akilleshäl: GFW blockerar regelmässigt dessa servrar men Tor experimenterar med flyktiga bryggservrar som inte är lika lätta att blockera [37]. För ändamålet används olika tekniker såsom CAPTCHAS och förtroendegrafer men trots detta tog det GFW en månad att blockera alla noder som tillhandahölls av dessa servrar.

GFW attackerar inte bara biblioteksservrar utan använder också Deep Packet Inspection (DPI) för att leta rätt på pågående nätverkstrafik mellan klienter och Tor-noder. Så fort en misstänkt nod hittas låtsas GFW vara en klient och påbörjar en protokollhandskakning med den misstänkta noden. Om det visar sig vara en nod så blockeras dess IP-adress i 12 timmar. Efter ytterligare 12 timmar gör GFW en ny handskakning för att utröna om noden fortfarande är aktiv etc.

Inom ramen för denna kapprustning har Tor även introducerat pluggbara transportmekanismer [38] som gömmer Tors trafik för att den ska se ut som t ex vitt brus (Obfs4) eller som andra valida protokollflöden (Meek). Dessa alternativa transportmekanismer har varit framgångsrika men GFW ser ofta igenom även även dessa obfuskeringar: Vitt brus upptäcks med hjälp av entropikänsliga DPI-filter och att härma verkliga protokollflöden har visat sig svårt.

2 Mixnätverkens akilleshälar och alternativen

Kapprustningen mellan censorer och mixnätverk i allmänhet, och Tor i synnerhet, har i omgångar visat på sårbarheter som gör dem lätta att blockera: Tors val att erbjuda låg latens har tvingar gjort detta extra tydligt [39]. Av den anledningen presenteras en aldrig sinande ström av alternativa låglatenta mixnätverk [4] för att lösa detta problem. Varje nytt mixnätverk introducerar mera avancerade teorier. Ett exempel är Loopix som stratifierar noder med hjälp av Poisson-distribution i akt och mening att stärka graden av anonymitet; kanske är det en bra idé - kanske inte. Loopix finns fortfarande bara i den artikel som den initialt beskrevs i.

Exempel som helt överger tanken på mixnätverk finns också. Private Information Retrieval (PRI) tekniker [41] har föreslagits som ett sätt att dölja vilka mail som tillhör vilken mottagare. Även anonym epidemisk routing i ad-hoc nätverk [46][47][50] har föreslagits.

2.1 Private Information Retrieval - The Pynchon Gate

Författarna till ”The Pynchon Gate” [40] introducerade en anonym remailer som påstods erbjuda större motståndskraft mot global trafikanalys än vad Mixmaster och Mixminion förmådde. För att lyckas med detta använde de PRI-tekniker som dolde vilka mail som tillhörde vilken mottagare.

PRI har sina rötter i databasvärlden och används normalt för att dölja vilka tabeller en databasklient läser. Författarna hävdade att PRI är en sundare teori att basera anonymitet på än mixnätverk och därmed mindre sårbar för attacker [42]. Pynchon Gate förväntade sig dock att mail skickas med hjälp av Mixmaster/Mixminion remailers (som inte längre finns) och att PRI-tekniker endast användes för hämtning av mail. I [42] påstår dock författarna att det borde vara möjligt att använda PRI-writing [43] för att skicka mail, och på så sätt helt undvika mixnätverk.

Ett problem med Pynchon Gate var dock att den krävde en central nym-server som förväntades ta hand om all inkommande mail och PRI-hantering. Ingen färdig implementation finns heller men värt att notera är att upphovsmännen är Len Sassaman och Nick Mathewson; förgrundsfigurerna bakom både Mixmaster, Mixminion och Tor.

2.2 Anonym epidemisk routing i ad-hoc nätverk

En annan väg framåt skulle kunna vara att helt överge tanken på anonymitet i hierarkiska nätverk och i stället vända blicken mot mobila icke-hierakiska ad-hoc nätverk. Dessa nätverk förlitar sig inte alls på det ordinarie hierarkiska nätet utan låter kommunikation flöda från mobil enhet till mobil enhet tills destinaltion nås. Dessa opportunistiska nätverk har dock varit en pipe dream i mer än tio år [44] och de har fortfarande inte fått allmän spridning och ingen killer app är i sikte.

Routing i epidemisk nätverk baseras på store-and-forward tekniker och routingprotokoll anpassade för delay-toleranta nätverk och är väl utforskade [45][48]. Många optimerade routingprotokoll har föreslagits, t ex Spray-and-Wait [49] har föreslagits och den grad av anonymitet som kan erbjudas i studerats ingående i ett flertal undersökningar [46][47][50].

Studier har också visat att det går att hålla nere strömåtgången hos de mobila enheter som kopplar upp sig till epidemiska nätverk [50]; speciellt om hög latens är acceptabelt samt att alla tillgängliga kommunikationskanaler används, dvs inte bara WiFi utan även 5G-D2D, Bluetooth och NFC.

3 Mail as the killer app ¯_(ツ)_/¯

Det känns fruktlöst att tro sig kunna presentera ett anonymitetsnätverk som är signifikant mycket bättre än Tor, I2P och Freenet.

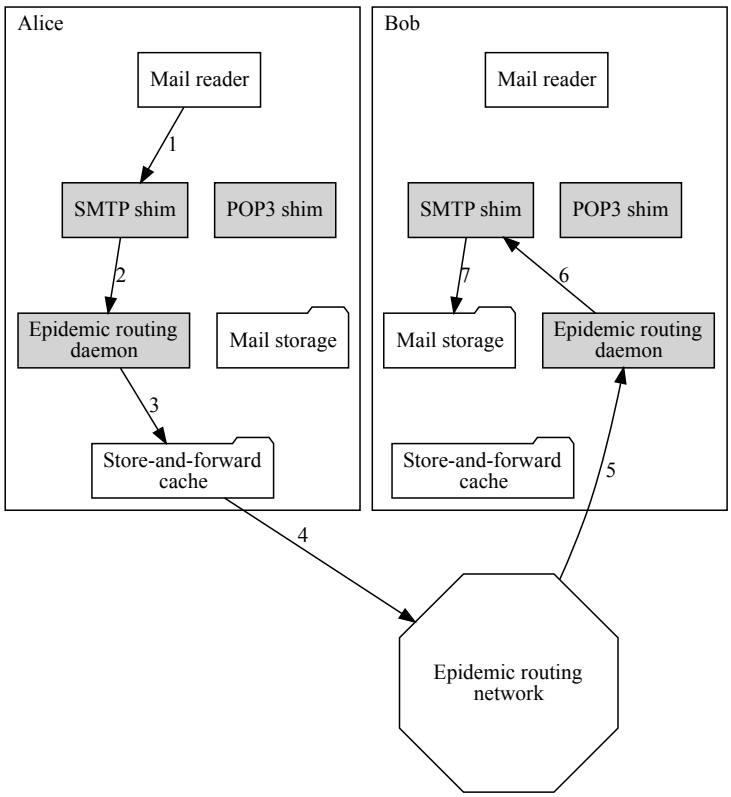
3.1 Ett blygsamt förslag

Anonym epidemisk routing i ad-hoc nätverk lämpar sig dock väl för just asynkrona tjänster såsom mail med sin tolerans för hög latens.

Ett naturligt sätt att göra en anonym mail-tjänst lättillgänglig skulle kunna vara att skriva en app bestående av SMTP/POP3 proxy shims [57] som lyssnar på localhost:25/110 och som under täcket realiserar ett anonymt epidemiskt ad-hoc nätverk. Användaren av den mobila enheten behöver bara installera en shims-app och skapa ett konto i mail-läsaren som pekar på SMTP/POP3 shimsen för att få tillgång till en anonymiserad mail-tjänst.

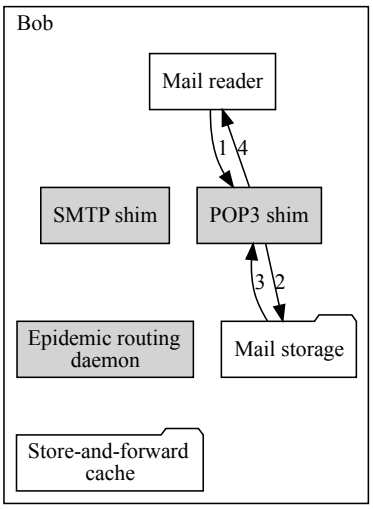
Denna insikt är inte på något sätt unik och finns beskriven på konceptnivå i t ex ”Opportunistic Email Distribution and Access in Challenged Heterogeneous Environments” [19] men jag noterar att ingen sådan tjänst ännu lanserats.

På schematisk nivå skulle det kunna se ut som i figur 1 där Alice skickar ett anonymt mail till Bob.



Figur 1: Alice skickar ett mail till Bob över det epidemiska nätverket

I figur 2 hämtar Bob mailet.

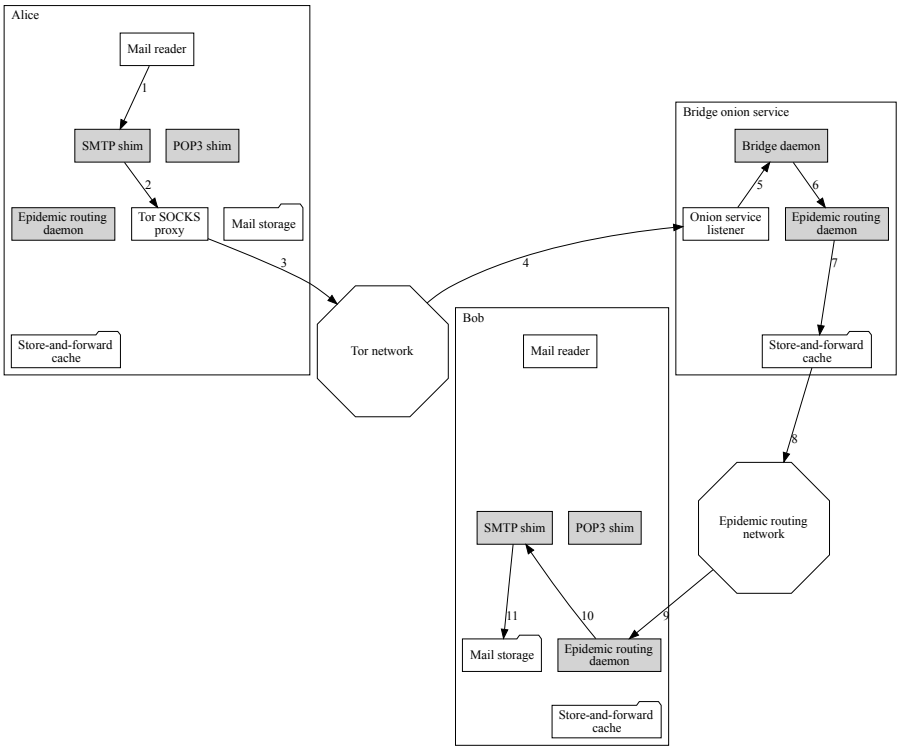


Figur 2: Bob läser mailet från Alice

Epidemiska ad-hoc nätverk är väl lämpade för mail-kommunikation på lokal nivå men på regional och global nivå blir latensen snabbt för hög (dagar) och en delmängd av alla skickade mail kommer inte att nå sina mottagare.

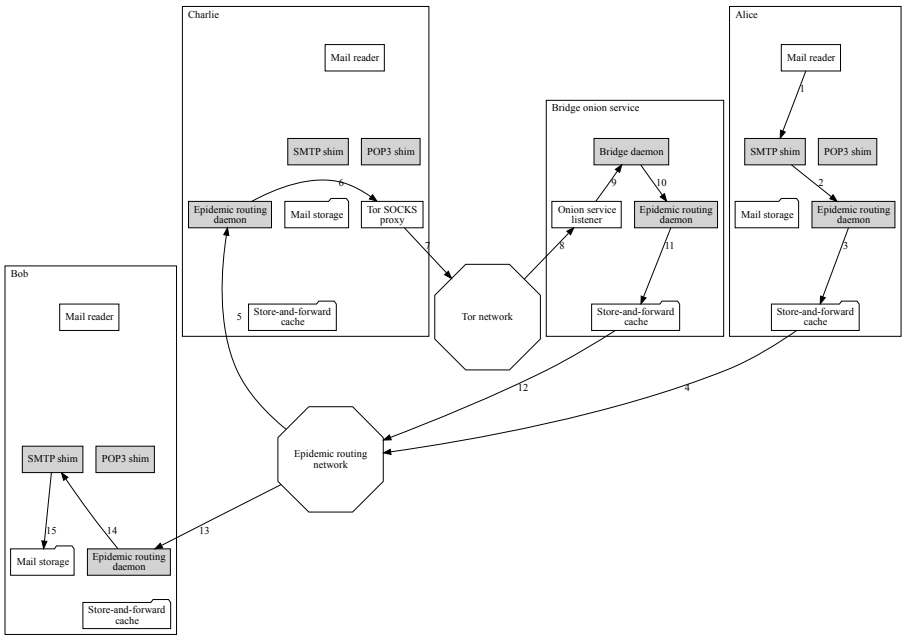
Antag dock att ett fåtal procent av de som installerat shims-appen också installerat Tors Orbot proxy-app [51] som ger tillgång till Tor-nätverket]. Även om bara några procent gör detta så kan dessa enheter användas som bryggor mellan olika delar av det epidemiska nätverket. På detta sätt kan utgående mail smita ut på Tor-nätverket i stället för att sakta ta sig fram via ren epidemisk routing.

I figur 3 har Alice valt att installera Orbot och när hon skickar ett mail till Bob smiter det direkt ut på Tor via Orbot och landar hos en Onion Service som bryggar mail vidare ut på den del av det epidemiska nätverket som just den Onion Servicen har kontakt med.



Figur 3: Alice skickar ett mail till Bob över sin Tor SOCKS proxy, via en Bridge Daemon som går som en Onion Service, och ut på det epidemiska nätverket igen, för att till slut anlända hos Bob

I figur 4 har inte Alice Orbot installerat men det har däremot Charlie. När Alice skickar ett mail till Bob så sänds det ut på det epidemiska nätverket men när det når Charlie så smiter det ut på Tor precis som i figur 3.



Figur 4: Alice skickar ett mail till Bob över det epidemiska nätverket, via Charlies Tor SOCKS proxy och en Bridge Daemon som kör som en Onion Service, och ut på det epidemiska nätverket igen, för att till slut anlända hos Bob

You get the gist.

3.2 Logisk grund

Skälen till att använda mail som killer app för att introducera stark anonymitet och censurmotstånd är många. Till att börja med är användare av mail vana vid att det kan ta tid för mail att komma fram. Användare behöver inte heller lära sig en ny app/metafor för att få tillgång till en anonymiserad mail-tjänst. De behöver bara installera shimsen och lägga till ett mail-konto i mail-läsaren. Eventuellt så läggs mail-kontot till automatiskt vid installation av shims-appen.

En killer app är förstås bara en killer app om det finns ett tillräckligt stort behov. Dissidenter i repressiva stater har uppenbarligen behov av just stark anonymitet och härdighet mot censur. Kanske borde fokus ligga på att erbjuda anonyma mail till just den gruppen och resten får följa i sinom tid: Inte bara av marknadsföringsmässiga skäl utan också för att epidemisk routing kommer till ett pris, dvs potentiellt hög latens och i värsta fall mail som aldrig levereras. (sic!)

3.3 Nedsidor och uppsidor

Anonym epidemisk routing såsom det beskrivs i t ex [46][47][50] har nersidor i termer av kommunikationskvalité men den är dock svår att blockera för en censor då den inte förlitar sig på centrala nätverkskomponenter. Förmågan att hantera partiell/hel nedtagning av nät, såsom gjordes under den arabiska våren, är en också en stark uppsida.

Att just kombinera anonym epidemisk routing med bryggning via Tor utnyttjar uppsidor hos båda dessa tekniker och gör det möjligt att hantera många olika typer attacker mot anonymitet och försök till censur.

3.4 Utökad funktionalitet

Kanske kan egenskaperna hos mail-transporten också styras av den som skickar mail med hjälp av plus-notation [53]. Några exempel:

`bob+skiptor@obscrete.org` - Skicka ett mail till Bob men brygga aldrig över till Tor (ens om det går)

`bob+skiptor;$arabspring42$libyafreedom@obscrete.org` - Skicka ett mail till Bob men brygga aldrig över till Tor (ens om det går) och använd bara noder i det epidemiska nätverket som märkts med *arabspring* och *libyafreedom*.

`bob+aging=4h@obscrete.org` - Skicka ett mail till Bob men låt aldrig mailet ligga kvar längre än fyra timmar på någon nod i det epidemiska nätverket.

Det är inte heller orimligt att tro att en anonymiserad mail-tjänst bara är ett skohorn in i det som skulle kunna kallas stark dissidentanonymitet: Dissidenter behöver rimligen i förlängningen också tjänster för micro-bloggning, anslagstavlor och diskussionsforum etc som också har dessa egenskaper.

3.5 Hur?

Hur skulle ett sådant epidemiskt nätverk realiseras? Vilket anonymt epidemiskt ad-hoc routingprotokoll är lämpligt att använda? Vilken grad av anonymitet och censurmotstånd kan uppnås? Detta kräver vidare utforskningar som inte ryms i detta dokument men god vägledning finns i t ex [46][47][50].

3.6 En prototyp

Jag har valt att i ett separat dokument med titeln ”A protoype: Censorship-resistant high-latency mail communication” [59] föreslå vilka tekniker som kan ligga till grund för en prototyp som realiserar det ovan beskrivna epidemiska nätverket, med tillhörande shims och Tor-bryggor.

Tanken är ett levande dokument som beskriver de implementationssteg som behövs för att bygga en fungerande prototyp med fokus på ”vad” och ”hur”, med tillhörande tidsestimat för varje steg.

Referenser

[1] https://en.wikipedia.org/wiki/Great_Firewall
[2] https://en.wikipedia.org/wiki/National_Security_Agency
[3] https://github.com/joagre/anond
[4] Free Haven's Selected Papers in Anonymity; https://www.freehaven.net/anonbib/
[5] Selected Research Papers in Internet Censorship; https://censorbib.nymity.ch
[6] https://www.torproject.org
[7] https://nymtech.net/
[8] https://www.coindesk.com/nym-technologies-raises-2-5-million-to-anonymize-crypto-apps
[9] The Loopix Anonymity System; Ania M. Piotrowska, Jamie Hayes, Tariq Elahi, Sebastian Meiser och George Danezis; https://www.usenix.org/system/files/conference/usenixsecurity17/sec17-piotrowska.pdf
[10] https://en.wikipedia.org/wiki/Mix_network
[11] Low Latency High Bandwidth Anonymous Overlay Network with Anonymous Routing; Roman Schlegel och Duncan S. Wong; https://github.com/joagre/anond/blob/master/doc/Schlegel-Wong-1.pdf
[12] Monotonically Increasing Bit Vector for Authenticated Anonymous Routing; Roman Schlegel and Duncan S. Wong; https://github.com/joagre/anond/blob/master/doc/Schlegel-Wong-2.pdf
[13] Anonymous overlay network supporting authenticated routing; Roman Schlegel and Duncan S. Wong; https://github.com/joagre/anond/blob/master/doc/Schlegel-Wong-3.pdf
[14] https://en.wikipedia.org/wiki/Tor_(anonymity_network)
[15] A Survey on Routing in Anonymous Communication Protocols; Fatemeh Shirazi, Milivoj Simeonovski, Muhammad Rizwan Asghar and Michael Backes; https://www.esat.kuleuven.be/cosic/publications/article-2706.pdf
[16] Threat modeling and circumvention of Internet censorship; David Fifield; http://digitalassets.lib.berkeley.edu/etd/ucb/text/Fifield_berkeley_0028E_17674.pdf
[17] https://en.wikipedia.org/wiki/Darknet
[18] Tor: Onion Service Protocol; https://2019.www.torproject.org/docs/onion-services.html.en
[19] Opportunistic Email Distribution and Access in Challenged Heterogeneous Environments; Tuomo Hyyryläinen; http://www.netlab.tkk.fi/~jo/dtn/2007-chants-dtn-mail.pdf
[20] Tor Metrics: Traffic; https://metrics.torproject.org/bandwidth-flags.html
[21] Tor Metrics: Onion Services; https://metrics.torproject.org/hidserv-rend-relayed-cells.html
[22] Anonymity Trilemma: Strong Anonymity, LowBandwidth Overhead, Low Latency—Choose Two; Debajyoti Das, Sebastian Meiser, Esfandiar Mohammadi och Aniket Kate; https://eprint.iacr.org/2017/954.pdf
[27] https://freenetproject.org/
[28] https://geti2p.net/en/
[29] https://en.wikipedia.org/wiki/Anonymizer
[30] https://www.silentcircle.com/

[31] https://en.wikipedia.org/wiki/Anonymous_remailer
[32] https://en.wikipedia.org/wiki/Cypherpunk_anonymous_remailer
[33] https://en.wikipedia.org/wiki/Mixmaster_anonymous_remailer
[34] https://en.wikipedia.org/wiki/Mixminion
[35] https://en.wikipedia.org/wiki/Pseudonymous_remailer
[37] Design of a blocking-resistant anonymity systemTor Project technical report; Roger Dingledinem, Nick Mathewson; https://svn.torproject.org/svn/projects/design-paper/blocking.pdf
[38] https://2019.www.torproject.org/docs/pluggable-transports
[39] Tor Networking Vulnerabilities and Breaches; Niketan Pate; http://www.cs.tufts.edu/comp/116/archive/fall2016/npatel.pdf
[40] The Pynchon Gate; Len Sassaman, Bram Cohen, Nick Mathewson; https://www.freehaven.net/anonbib/cache/sassaman:wpes2005.pdf
[41] https://en.wikipedia.org/wiki/Private_information_retrieval
[42] https://youtu.be/lb9JTaneySs
[43] Distributed Point Functions and their Applications; Niv Gilboa1 och Yuval Ishai2; https://www.iacr.org/archive/eurocrypt2014/84410245/84410245.pdf
[44] A Decade of Research in Opportunistic Networks – Challenges, Relevance, and Future Directions; Sacha Trifunovic, Sylvia T. Kouyoumdjieva, Bernhard Distl, Ljubica Pajevic, Gunnar Karlsson, Bernhard Plattner; https://people.kth.se/~stkou/pub/commag2017.pdf
[45] https://en.wikipedia.org/wiki/Routing_in_delay-tolerant_networking
[46] Toward Delay Tolerant Network Anonymity; Rob Jansen och Robert Beverly; https://pdfs.semanticscholar.org/d4b8/ea63f9be49bbd075531fe116a88c5a3419f4.pdf
[47] Anonymity and Security inDelay Tolerant Networks; Aniket Kate, Gregory M. Zaverucha, Urs Hengartner och David R. Cheriton; http://www.cs.bham.ac.uk/~tpc/cwi/Teaching/MASPPapers/RAN.pdf
[48] Epidemic Routing for Partially-Connected Ad Hoc Networks; Amin Vahdat och David Becke, http://issg.cs.duke.edu/epidemic/epidemic.pdf
[49] Spray and Wait: An Efficient Routing Scheme for Intermittently Connected Mobile Networks; Thrasyvoulos Spyropoulos, Konstantinos Psounis och Cauligi S. Raghavendra; http://chants.cs.ucsb.edu/2005/papers/paper-SpyPso.pdf
[50] Performance and Energy Consumption Analysis of a Delay-Tolerant Network for Censorship-Resistant Communication; Yue Liu, David R. Bild, David Adrian, Gulshan Singh, Robert P. Dick, Dan S. Wallach och Z. Morley Mao; http://robertdick.org/publications/liu15jun.pdf
[51] https://guardianproject.info/apps/orbot/
[52] https://www.torproject.org/download/
[53] https://www.cs.rutgers.edu/~watrous/plus-signs-in-email-addresses.html
[54] Untraceble Electronic, Return Addresses, and Digital Pseudonyms; David L. Chaum; https://www.freehaven.net/anonbib/cache/chaum-mix.pdf
[55] https://sv.wikipedia.org/wiki/Onion_routing
[56] https://www.torproject.org/about/sponsors/
[57] https://en.wikipedia.org/wiki/Shim_(computing)
[58] https://en.wikipedia.org/wiki/SOCKS
[59] A protoype: Censorship-resistant high-latency mail communication; Joakim G.