

Studi Banding Spam-Assassin Mail Server Dengan dan Tanpa Filter di Sisi Mail Client

Nurlina^{*1}, Irmayana²

¹Sistem Informasi STMIK DIPANEGARA Makassar

²Teknik Informatika STMIK DIPANEGARA Makassar

E-mail: *¹linanoer888@yahoo.com, ²irmayana@yahoo.com

Abstrak

Filter spam yang disediakan oleh situs penyedia layanan email seperti yahoo, gmail, aim mail, windows live hotmail, dan masih banyak lagi yang lainnya, merupakan fasilitas yang sangat bermanfaat bagi para user-nya. Filter spam tidak akan berfungsi sebagaimana yang diharapkan oleh user client. Pada saat alamat email para client sudah pernah di subscribe dengan tujuan tertentu seperti misalnya untuk registrasi mailing list, newsgroup, dan lain sebagainya, maka alamat emailnya itu sudah tidak aman lagi dari para spammer. Pada dasarnya para admin mail server hanya menggunakan secara langsung filter spam yang disediakan oleh mail server yang diinstal, tanpa memberikan penyettingan tertentu yang dibutuhkan client sama sekali. Para user sendiri yang seharusnya lebih aktif dalam menyaring spam pada email yang digunakan dengan banyak cara. Penelitian ini memanfaatkan aplikasi mail client Thunderbird untuk menyaring spam dengan metode Bayesian sebagai kelanjutan dari menyaring spam yang sudah tersaring sebelumnya pada sisi mail server dan bertujuan menganalisis hasil pengklasifikasian email ham dan email spam pada mail server dan mail client. Disimpulkan bahwa nilai akurasi dan error filter spam pada mail server berhubungan dengan filter Spam-Assassin yang tidak disetting dan dikonfigurasi oleh adminnya menunjukkan hasil yang tidak memuaskan dibandingkan dengan filter spam metode bayesian pada mail client yang sudah di-training.

Kata Kunci — Studi Banding, Spam Assassin, Mail Server, Metode Bayesian, Mail Client

Abstract

Spam filters provided by your email service provider websites such as yahoo, gmail, AIM mail, windows live hotmail, and many others, is a very powerful feature for the user. The spam filter will not work as expected by the client user. at the time of the email address of the client has been ever subscribe to a specific purpose such as for registration, mailing lists, newsgroups, and so forth, then the email address is no longer safe from spammers. Basically the admin mail server directly using only the spam filter provided by the mail server is installed, without giving a specific setting it takes the client at all. The users themselves are supposed to be more active in the spam filter on the email that is used in many ways. This study utilizes Thunderbird mail client application to filter spam with Bayesian methods as a continuation of the spam filter that has been previously filtered on the mail server and to analyze the results of the classification of ham and spam e-mail on the mail server and mail client. It was concluded that the accuracy and error spam filter on the mail server associated with the filter Spam-Assassin is not be set and configured by the admin showed unsatisfactory results compared with the Bayesian method to filter spam mail client that is already in-training

Keywords — Comparative Study, Spam-Assassin, Mail Server, Metode Bayesian, Mail Client

1. PENDAHULUAN

Penelitian ini dilakukan awalnya karena peneliti selalu mendapatkan email spam setiap hari pada emailnya dengan domain yahoo.com. Seperti kita ketahui, domain yahoo.com sudah menyediakan folder khusus untuk email spam, namun kenyataannya tidak semua email spam langsung masuk pada folder yang sudah disediakan. Sedangkan email yang seharusnya masuk di inbox penerima justru masuk pada folder spam sehingga email yang penting terkadang tidak dibaca oleh yang bersangkutan. Peneliti menyelidiki email dengan domain yahoo.com menyediakan folder spam yang memanfaatkan anti spam yang bernama program Spam-Assassin. Anti spam ini merupakan program yang terintegrasi dari software mail server yang digunakan. Jadi secara otomatis penggunaannya juga merupakan *default* dari yang sudah disediakan oleh mail server. Spam-Assassin memiliki berbagai macam aturan menyangkut spam. Berangkat dari hal tersebut, peneliti mencoba mengangkat permasalahan ini yang menurut hematnya belum ada yang menyelidiki tentang aturan-aturan dari yang disediakan Spam-Assassin tersebut.

2. METODE PENELITIAN

Studi banding ini menggunakan dua metode yaitu metode *Spam-Assassin* dan metode *Bayesian*. Untuk mendapatkan data di lapangan digunakan teknik observasi dan wawancara. Menghimpun bahan-bahan keterangan yang dilakukan dengan mengadakan pengamatan dan pencatatan secara sistematis terhadap fenomena-fenomena yang dijadikan objek pengamatan dan dengan menghimpun bahan-bahan keterangan dengan tanya jawab secara lisan, sepihak, berhadapan muka dengan arah tujuan yang telah ditentukan.

Populasi dalam penelitian ini adalah keseluruhan email spam dan email ham yang masuk dalam inbox *account* linanoer888@unhas.ac.id yang jumlahnya dapat dilihat pada Tabel 1. Email spam juga berasal dari *account* linanoer888@yahoo.com, yang terakumulasi pada Tabel 2. Email ham yang berasal juga dari linanoer@unhas.ac.id seperti pada Tabel 3. Email-email yang tersimpan dalam *mailbox / maildir* server Universitas Hasanuddin.

Tabel 1. Populasi *account* linanoer888@unhas.ac.id

Total Emails	Email Ham	Email Spam
722	210	512
100%	29,09%	70,91%

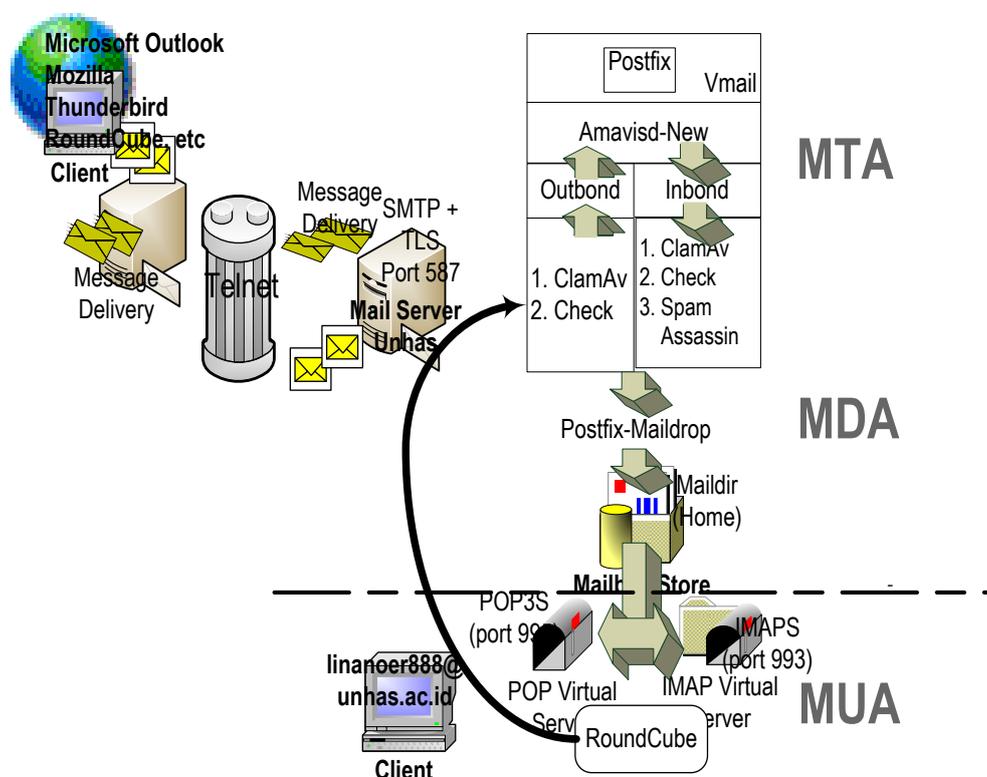
Tabel 2. Populasi *account* linanoer888@yahoo.com

Total Emails	Email Ham	Email Spam
274	18	256
100%	6,57%	93,43%

Tabel 3. Populasi *account* linanoer@unhas.ac.id

Total Emails	Email Ham	Email Spam
10	10	0
100%	100%	0%

Spesifikasi hardware untuk mail server sangat tergantung pada tingkat kebutuhan dan software yang digunakan. Kebutuhan hardware minimum untuk Apache sebagai mail server adalah komputer Intel ® Xeon CPU E540 @ 2.00 GB dengan RAM 1 GB dan HDD 250 GB. Struktur mail server yang digunakan di Universitas Hasanuddin seperti terlihat pada gambar 1, yang terdiri dari komponen-komponen: Apache sebagai Web Server, Linux Debian sebagai Sistem Operasi, Postfix sebagai MTA, Vmail untuk fungsi support Multi Domain pada MTA, Postfix-maildrop sebagai MDA untuk melakukan email scanning dan mengintegrasikan ClamAv serta Spam Assassin, Database MySQL, Amavisd-New sebagai Content-Checks antara ClamAv dan Spam Assassin, Spam-Assassin 3.2.5 sebagai Anti Spam, ClamAV sebagai Anti Virus, SASL + TLS/SSL sebagai Secure Authentication dan Encryption, belum memiliki fitur mailing_list, Courier IMAPS/POP3S/Devecot sebagai IMAP Server(IMAPS)/POP3S untuk menyediakan akses webmail melalui IMAPS (port 993)/POP3S (port 995), dan RoundCube sebagai MUA berbasis Webmail.



Gambar 1. Struktur Mail Server Universitas Hasanuddin
(Sumber: Pratama, 2008)

Dari gambar 3, tampak pemisahan sisi server (MTA & MDA) dari sisi client (MUA). Pada gambar tersebut juga ditunjukkan bagaimana hubungan interaksi antara MTA, MDA, dan MUA, serta keterlibatan komponen-komponen yang mendukung, baik yang berada di sisi MTA (ClamAV dan Spam-Assassin), di sisi MDA (Postfix-Maildrop), dan di sisi MUA (RoundCube). Pengiriman E-Mail ke internet dilakukan melalui SMTP Relay pada server SMTP milik ISP Telnet. Protokol SMTP berfungsi untuk mendownload pesan dari klien E-Mail ke server SMTP yang ada di ISP. Selanjutnya server SMTP melakukan pengiriman E-Mail ke alamat tujuan masing-masing. Karena tugas sebagai mediator tersebut maka server SMTP di ISP disebut sebagai SMTP relay. Server SMTP dari ISP Telnet memungkinkan untuk menerima relay E-Mail dari domain unhas.ac.id. Admin melakukan autentifikasi dengan user name dan password tiap user sebelum mengirim melalui server SMTP ISP Telnet. Hal ini dilakukan untuk mencegah tindakan spam yang sangat merugikan ISP maupun penerima E-Mail [6].

Pengambilan E-Mail di Unhas dilakukan dengan metode MultiPOP yakni menggunakan protokol POP3S/IMAPS untuk mendownload E-Mail dari server POP3S/IMAPS ke mail server jaringan. Sebelum mengambil E-Mail harus diberikan informasi mengenai account *user* serta password sehingga server POP3S/IMAPS dapat mengenali identitas mailbox yang akan didownload. E-Mail tiap *user* di simpan di mailboxnya masing-masing, tidak digabungkan dalam satu mailbox. Tiap account memiliki alamat dan password masing-masing sehingga harus dilakukan konfigurasi MultiPOP untuk tiap *user* di mail server. Mail Server berjalan dengan beberapa protokol pada TCP/IP, yakni SMTP dan TLS (port 587), POP3S (port 995), dan IMAPS (port 993). Mail Server memiliki tiga komponen utama yang menyusunnya, yakni Mail Transfer Agent (MTA), Mail Delivery Agent (MDA), dan Mail User Agent (MUA). MTA bertugas mengatur pengiriman dan penerimaan e-mail, MDA bertugas mengatur pengiriman e-mail ke alamat yang sesuai pada jaringan lokal, sementara MUA bertugas untuk menjadi antarmuka yang menghubungkan *user* dengan Mail Server [6].

Studi banding (*comparison study*) biasa dilakukan untuk maksud peningkatan mutu, perluasan usaha, perbaikan sistem, penentuan kebijakan baru, perbaikan peraturan perundangan, dan lain-lan. Kegiatan studi banding dilakukan oleh kelompok kepentingan untuk mengunjungi atau menemui obyek tertentu yang sudah disiapkan dan berlangsung dalam waktu relatif singkat. Intinya adalah untuk membandingkan kondisi obyek studi di tempat lain dengan kondisi yang ada di tempat sendiri. Hasilnya berupa kumpulan data dan informasi sebagai bahan acuan dalam perumusan konsep yang diinginkan[1].

Spam-Assassin adalah aplikasi yang sudah teruji secara luas menggunakan proyek open source yang berfungsi sebagai mail filter untuk mengidentifikasi spam. *Spam-Assassin* berjalan pada server, dan sebagai filter spam sebelum mencapai kotak surat *user*. *Spam-Assassin* menyaring *e-mail* masuk sehingga *e-mail* account milik pelanggan web hosting terlindungi dari SPAM. *Spam-Assassin* diintegrasikan ke dalam mail server untuk secara otomatis menyaring semua spam untuk domain dan aturan penggunaan, atau "tes," untuk menentukan apa yang spam dan apa yang tidak. *Spam-Assassin* dapat memberikan tanda dengan mengubah subject *e-mail* atau langsung menghapus *e-mail* SPAM yang masuk. Setiap *user* dapat mengkonfigurasi ambang batas *Spam-Assassin* untuk menentukan jika sebuah email adalah spam, dicurigai spam, atau bukan spam. Spam-Assassin dirancang untuk memberikan penyaringan yang dapat diterima tanpa pelatihan apapun dan tentu saja bisa diperbaiki jika tiap user menghabiskan waktu untuk melatihnya [7].

Spam-Assassin menggunakan berbagai mekanisme berikut [7] :

1. Pengecekan *Header* Email
2. Pengecekan isi email
3. Pengelompokan Email *Address* secara otomatis/manual menjadi *Whitelist* ataupun *Blacklist*
4. Bayesian *Filtering*
5. Penyaringan database spam kolaboratif (DCC, Pyzor, Razor2)
6. Berbasis jaringan seperti *Blacklist* URL, *Blacklist* DNS, *Checksum* berbasis *Filter*, dan Algoritma *Hash*

Mail Server merupakan *server* yang melayani beberapa komputer dalam suatu jaringan internet, ekstranet dan intranet dalam bentuk layanan pengiriman dan pengambilan *e-mail*. Mesin ini senantiasa menerima pesan dari *email client* yang digunakan user atau mungkin dari *email server* yang lainnya. Sebuah *mail server* biasanya terdiri dari area penyimpanan, set konfigurasi *user*, daftar *user* dan seri modul komunikasi [2]. Mail Server adalah suatu entitas berupa komputer yang bertindak sebagai sebuah server (penyedia layanan) dalam jaringan komputer / internet, serta memiliki fungsi untuk melakukan penyimpanan (storing) dan distribusi yang berupa pengiriman (sending), penjaluran (routing), dan penerimaan (receiving) e-mail. Mail Server berjalan dengan beberapa protokol pada TCP/IP, yakni SMTP (port 25), POP3 (port 110), dan IMAP (port 143). Mail Server memiliki tiga komponen utama yang menyusunnya, yakni Mail Transfer Agent (MTA), Mail Delivery Agent (MDA), dan Mail User Agent (MUA). MTA bertugas mengatur pengiriman dan penerimaan e-mail, MDA bertugas

mengatur pengiriman e-mail ke alamat yang sesuai pada jaringan lokal, sementara MUA bertugas untuk menjadi antarmuka yang menghubungkan user dengan Mail Server [5].

a. *Storage Area*

Adalah tempat pesan *email* disimpan untuk *user* dan sebagai transit sebelum menuju tujuan lainnya.

b. *Set konfigurasi user*

Adalah untuk *user* yang menetapkan bagaimana *mail server* harus beraksi saat menerima pesan tertentu atau mungkin mengambil keputusan untuk pengirim tertentu. Sebagai contoh, *address* tertentu mungkin dibatasi untuk hanya dapat mengirim pesan dalam lingkungan perusahaan saja.

c. *Daftar/list user*

Adalah *database account user* yang dikenali *mail server* dan akan berkomunikasi dengan mereka secara lokal.

d. *Modul komunikasi*

Adalah komponen yang aktualnya menangani transfer pesan ke dan dari *mail server* lain dan atau klien lain. Modul yang terinstal dalam *mail server* ini dapat berbeda-beda sesuai kebutuhan.

Mail server hanya sebuah aplikasi yang berurusan dengan lalu lintas *email*. Tidak secara langsung berhubungan dengan *user* yang akan berkirim *email*. Dalam pengiriman *email*, terdapat dua aplikasi yang diperlukan yaitu MTA (*Mail Transfer Agent*), dan MUA (*Mail User Agent*).

Secara garis besar MTA adalah aplikasi *server* yang berfungsi untuk mengirimkan *e-mail* dari *mail server* local ke *mail server remote*. Salah satu contoh jenis MTA yang terkenal dan digunakan dalam proyek akhir ini adalah postfix [2].

Berikut beberapa fungsi dari MTA:

1. Pertukaran *email* menggunakan protocol TCP.
2. Menerima *email* masuk (*incoming*).
3. Meneruskan *email* yang akan keluar (*outgoing*).

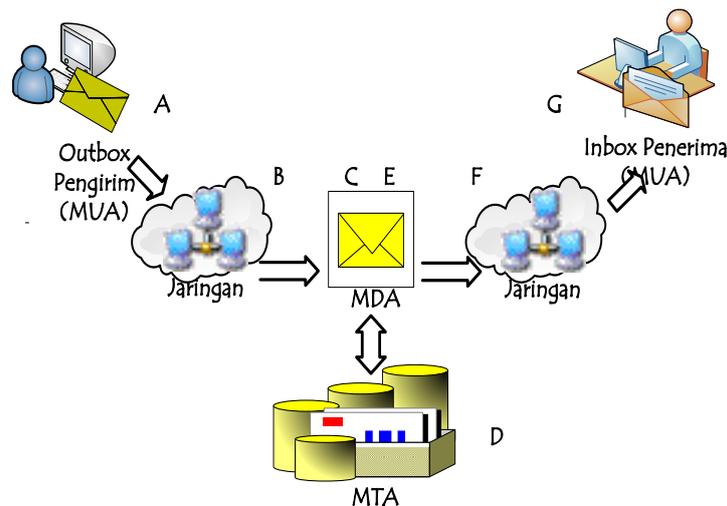
Sedangkan MUA adalah aplikasi yang berfungsi sebagai *interface* antara *email*, dalam hal ini berhubungan dengan *user* yang memiliki *email* tersebut, dengan MTA yang mendukungnya. Berfungsi sebagai berikut:

1. Menulis *email* dan membaca *email* yang masuk [2].
2. Mengatur konfigurasi *email* sehingga sesuai dengan MTA yang mendukungnya.

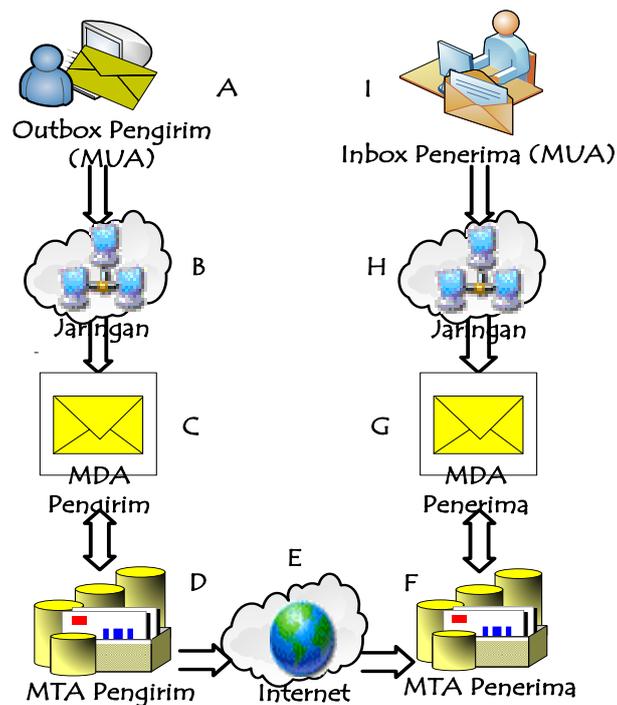
Secara ringkas, cara kerja Mail Server ditunjukkan dalam gambar 2 dan gambar 3 [5].

Email client, email pembaca, atau lebih formal disebut mail user agent (MUA), adalah sebuah program komputer yang digunakan untuk mengelola email. Secara khusus, istilah email client dapat merujuk kepada setiap agen yang bertindak sebagai client ke mail server, terlepas dari itu menjadi pengguna email agen, menyampaikan ke server, atau diketikkan manusia pada sebuah terminal. Selain itu, sebuah aplikasi web menyediakan pengelolaan pesan, komposisi, dan fungsi penerimaan yang kadangkadang dianggap sebagai email client.

Seperti kebanyakan program client, sebuah MUA hanya aktif ketika pengguna (*user*) menjalankannya. Pesan tiba di Mail Transfer Agent (MTA) server. Kecuali MUA yang memiliki akses ke penyimpanan server, pesan akan disimpan di sebuah server remote dan MUA harus mengambil mereka atas permintaan pengguna email.



Gambar 2. Proses Pengiriman Email pada Jaringan Lokal
(Sumber: Pratama, 2008)



Gambar 3. Proses Pengiriman Email ke Jaringan Luar
(Sumber: Pratama, 2008)

Bayesian filter atau **Naive Bayes Classifier** merupakan metode terbaru yang digunakan untuk mengklasifikasikan sekumpulan dokumen. Algoritma ini memanfaatkan metode probabilitas dan statistik yang dikemukakan oleh ilmuwan Inggris Thomas Bayes, yaitu memprediksi probabilitas di masa depan berdasarkan pengalaman di masa sebelumnya. Dua kelompok peneliti, satu oleh Pantel dan Lin, dan yang lain oleh Microsoft Research memperkenalkan metode statistik Bayesian ini pada teknologi anti spam filter. Tetapi yang membuat algoritma Bayesian filtering ini popular adalah pendekatan yang dilakukan oleh Paul Graham [6].

Dasar dari teorema naive Bayes [6], yang digunakan dalam pemrograman adalah rumus Bayes berikut ini:

$$P(A|B) = (P(B|A) * P(A))/P(B) \dots\dots\dots (1)$$

Peluang kejadian A sebagai B ditentukan dari peluang B saat A, peluang A, dan peluang B. Pada pengaplikasiannya nanti rumus ini berubah menjadi:

$$P(C_i|D) = (P(D|C_i)*P(C_i)) / P(D) \dots\dots\dots(2)$$

Naive Bayes Classifier atau bisa disebut sebagai multinomial naive bayes merupakan model penyederhanaan dari algoritma bayes yang cocok dalam pengklasifikasian text atau dokumen [8].

Persamaannya adalah:

$$v_{MAP} = \arg \max_{v_j \in V} P(v_j | a_1, a_2, \dots, a_n) \dots\dots\dots(3)$$

Menurut persamaan (3), maka persamaan (1) dapat ditulis

$$v_{MAP} = \arg \max_{v_j \in V} \frac{P(a_1, a_2, \dots, a_n | v_j)P(v_j)}{P(a_1, a_2, \dots, a_n)} \dots\dots(4)$$

$P(a_1, a_2, \dots, a_n)$ konstan, sehingga dapat dihilangkan menjadi

$$v_{MAP} = \arg \max_{v_j \in V} P(a_1, a_2, \dots, a_n | v_j)P(v_j) \dots\dots(5)$$

Karena $P(a_1, a_2, \dots, a_n | v_j)$ sulit untuk dihitung, maka akan diasumsikan bahwa setiap kata pada dokumen tidak mempunyai keterkaitan.

$$v_{MAP} = \arg \max_{v_j \in V} P(v_j) \prod_i P(a_i | v_j) \dots\dots(6)$$

Keterangan :

$$P(v_j) = \frac{| docs_j |}{| Contoh |} \dots\dots\dots(7)$$

$$P(w_k | v_j) = \frac{n_k + 1}{n + | kosakata |} \dots\dots\dots(8)$$

Dimana untuk :

- $P(v_j)$: Probabilitas setiap dokumen terhadap sekumpulan dokumen.
- $P(w_k|v_j)$: Probabilitas kemunculan kata w_k pada suatu dokumen dengan kategori class v_j
- $| docs |$: frekuensi dokumen pada setiap kategori
- $| Contoh |$: jumlah dokumen yang ada
- n_k : frekuensi kata ke-k pada setiap kategori
- $kosakata$: jumlah kata pada dokumen test

Pada persamaan (8) terdapat suatu penambahan 1 pada pembilang, hal ini dilakukan untuk mengantisipasi jika terdapat suatu kata pada dokumen uji yang tidak ada pada setiap dokumen data training [1].

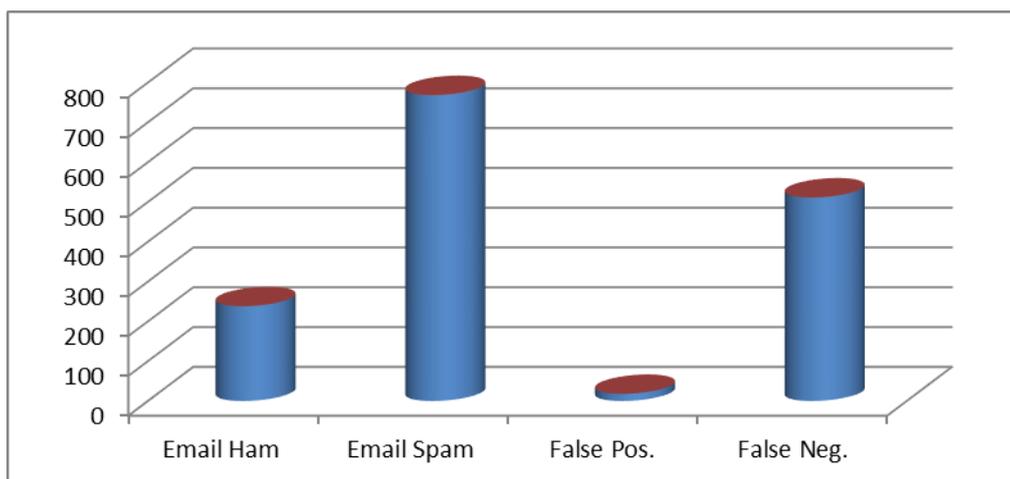
3. HASIL DAN PEMBAHASAN

Tabel 5 merupakan hasil analisa dan perhitungan data dengan mengamati secara umum data email ham yang dianggap spam disebut dengan *false positive* dan pengamatan data email spam yang lolos dari *filter*/anti spam yang disebut dengan *false negative*, langsung pada *inbox mail client* dengan account linanoer888@unhas.ac.id. Rasio antara email ham dan email spam tidak memberikan efek yang berarti pada *filter*, maksudnya bahwa rasio 1:1 juga tidak akan memberikan akurasi lebih baik dibanding rasio 5:1 [4].

Sedangkan pada sisi client, peneliti menggunakan aplikasi mail client *Thunderbird* pada *notebook* dengan sistem operasi *Microsoft Windows XP Professional Versi 2002 Service Pack 2*, Intel Atom 1.66 GHz dan RAM 1 GB. Penggunaan daripada metode Bayesian *Filter* dalam penelitian ini dengan menginstalasi dan mengkonfigurasi software *Thunderbird 1.0* atau *Thunderbird 2.0*. *Thunderbird* dipilih karena *user friendly*, *freeware* dan lintas *platform*. Pelatihan dan pengujian dilakukan sebanyak limabelas kali (15x) *training* dan total email yang digunakan dalam *training* sebanyak 996 email yang terdiri dari 233 email ham dan 768 email spam. Sedangkan data yang digunakan untuk mengujinya sebanyak 276 email spam dan 23 email ham. Tabel 6 merupakan hasil pengujian metode bayesian dengan menggunakan *Thunderbird* pada sisi client. Pengukuran kinerja dilihat dari *false positive* dan *false negative* atau *accuracy* dan *error* -nya.

Tabel 4. Hasil Analisa metode *Spam-Assassin* pada Mail Server Unhas

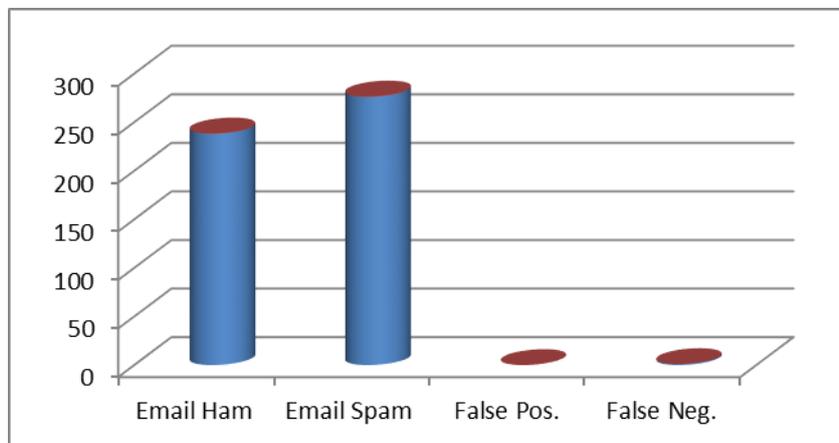
Total Emails	Email Ham	Email Spam	False Pos.	False Neg.	Acc	Err
1006	238	768	18	511		
100%	23.66%	76.34%	1.79%	50.80%	47.89%	52.11%



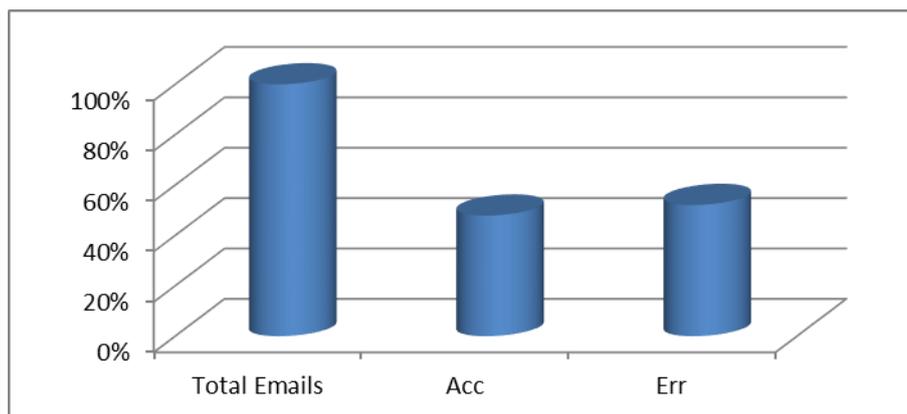
Gambar 4. Grafik Hasil Analisa Metode *Spam-Assassin* pada Mail Server Unhas

Tabel 5. Hasil Pengujian Metode Bayesian pada Mail Client *Thunderbird User*

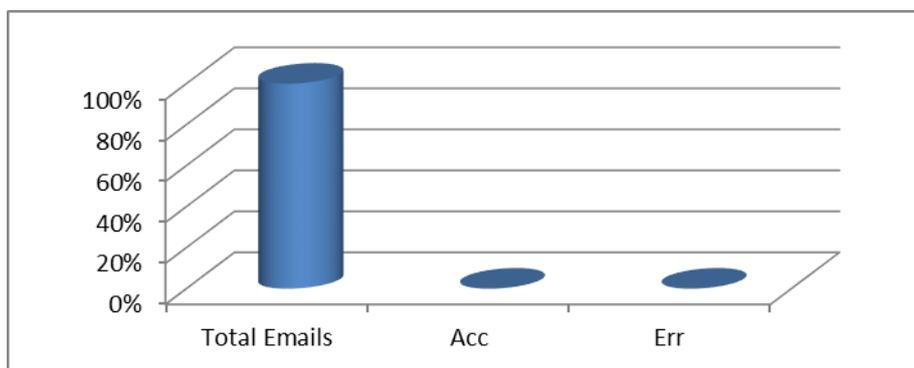
Total Emails	Email Ham	Email Spam	False Pos.	False Neg.	Acc	Err
514	238	276	0	1	99,98%	0,02%



Gambar 5. Grafik Hasil Pengujian Metode Bayesian pada Mail Client *Thunderbird User*



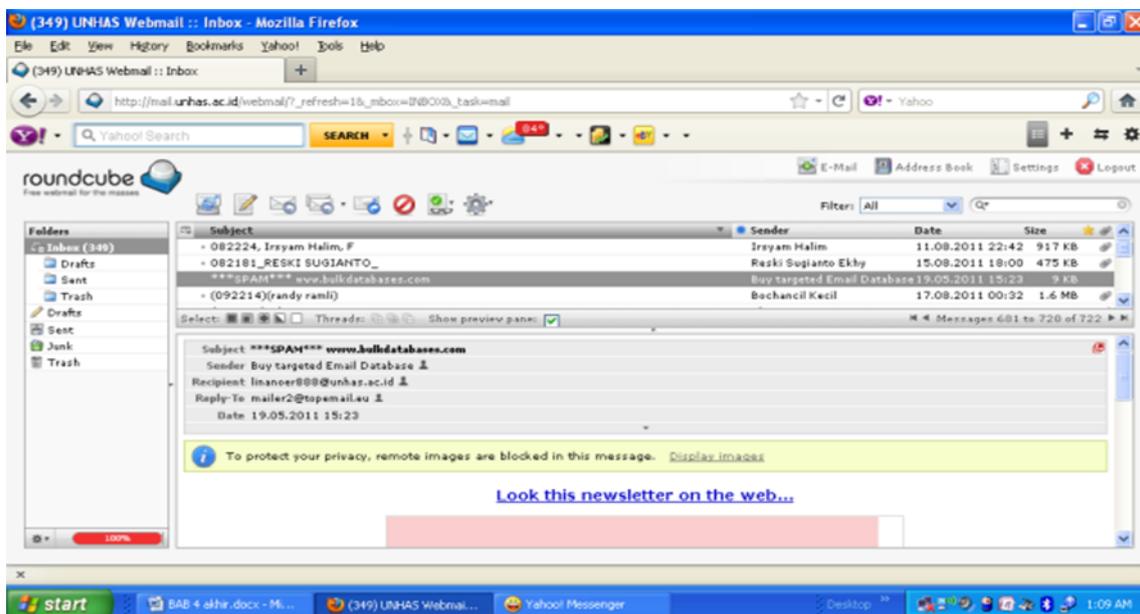
Gambar 6. Grafik Hasil Akurasi dan Error Metode *Spam-Assassin* pada Mail Server Unhas



Gambar 7. Grafik Hasil Akurasi dan Error Metode *Bayesian* pada Mail Client *Thunderbird*

Berdasarkan pengamatan dari Tabel 4 dan Tabel 5 menunjukkan bahwa penggunaan metode *Spam-Assassin* pada Mail Server Unhas mencapai nilai akurasi 47,39% dan nilai error 52,61%. *False Positive* mencapai angka 1,31% yakni ada 13 email ham yang dianggap spam

dan tidak masuk ke dalam inbox *account* peneliti. Hal ini sangat mengkhawatirkan karena email yang kita butuhkan informasinya menjadi tidak sampai dalam inbox kita. Beberapa hal yang menyebabkan demikian adalah berasal dari aturan *Spam-Assassin* 3.2.5 sendiri[8], yakni pengetesan *Spam-Assassin* 3.2.5 pada daerah '*header*', SUBJ_ALL_CAPS, seluruh *subject* berisi huruf besar (huruf *capital*) dengan 1,926 point, FROM_ILLEGAL_CHARS, asal email (*from* dan *subject*) dituliskan dengan banyaknya karakter yang ilegal dengan 3,999 point dan 1,954 point, HEAD_LONG, terlalu panjangnya pesan yang dituliskan pada *header* nya dengan 2,5 points, pada daerah '*body*', FUZZY_SOFTWARE, berisi kata-kata yang dianggap mengaburkan dalam spam) dengan 3,169 point dan UNWANTED_LANGUAGE_BODY, pesan ditulis dalam bahasa yang tidak diinginkan dengan 2,8 point. Standar dari *Spam-Assassin* agar email spam tidak masuk inbox ataupun tidak lolos dari *filter* adalah sebesar 5 point. Sedangkan *False Negative* mencapai nilai 51,31% dengan 511 email spam yang lolos dari *filter*/anti spam *Spam Assassin*. Hanya satu email yang diberi tanda *****SPAM***** oleh *Spam-Assassin* seperti terlihat pada gambar 8.



Gambar 8. Struktur Mail Server Universitas Hasanuddin

Mengapa email tersebut diberi tanda dan lolos dari *filter* sedangkan email yang sama sebelumnya lolos dan tanpa tanda *****SPAM*****? Karena sebelumnya sudah banyak email seperti gambar 4 yang telah masuk dalam inbox *user* yang bersangkutan. Sehingga mendapat nilai *score* sesuai dengan aturan *Spam-Assassin* dengan nama tesnya masing-masing, DATE_IN_FUTURE_06_12, dengan 2,136 point, FH_MSGID_01C70XXX, nomor ID dari pesan tersebut sering kelihatan sebelumnya dalam jumlah banyak dengan 2,757 point dan TVD_SPACED_SUBJECT_WORD3, pada *Subject* terdapat kata FW/RE dengan 2,276 point.

Sebaiknya email dengan domain unhas.ac.id tersebut selanjutnya menggunakan pengujian metode Bayesian pada Mail Client sehingga bisa mencapai nilai akurasi sampai 99,98% dan nilai error hanya 0,02%. *Spam-Assassin* yang sebenarnya memiliki metode bayesian yang terintegrasi di dalam mail server akan bermanfaat dan bisa berhasil menyaring sampai 99,98% akuratnya bila di sisi server, adminnya langsung mentraining email *user* masing-masing, namun akan menyita waktu yang sangat lama bagi admin dan hal ini tidak akan mungkin dilakukan oleh siapa saja yang menjadi admin karena fasilitas metode bayesian yang disediakan harus juga senantiasa di *update* agar mendapatkan hasil yang maksimal. Seperti diketahui, *spammer* akan selalu meningkatkan cara dan pengetahuannya dalam menembus email dari korbannya. Maka dapat disimpulkan bahwa metode Bayesian akan sangat efektif dan cukup

akurat dengan kinerja lebih baik apabila disepadankan dengan metode *Spam-Assassin* pada Mail Server khususnya pada mail server Universitas Hasanuddin.

4. KESIMPULAN

1. Klasifikasi metode *bayesian* dapat digunakan dengan baik dalam aplikasi *email filtering* dari serangan spam, pada *level client* untuk menyaring spam.
2. Semakin banyak jumlah data yang digunakan untuk *training* pada metode bayesian maka semakin tinggi keakuratannya.
3. Semakin sering dilakukan *learning* atau penambahan *knowledge* pada metode bayesian di sisi *client* maka akan semakin cepat mendapatkan klasifikasi dengan benar.
4. *Email filtering* dengan metode *bayesian* masih memiliki akurasi yang tinggi meskipun jumlah data *trainingnya* sedikit.
5. Pada sisi server, sebaiknya para user meminta langsung pada administrator mail servernya untuk men-*setting score* dari Spam-Assassin sesuai dengan kebutuhan dari inboxnya.
6. Para user pengguna mail server sebaiknya melatih/ mentraining sendiri emailnya karena hal ini bukanlah kewajiban dari para administrator

5. SARAN

1. Di sisi *client* aplikasi dapat dikembangkan lebih lanjut, misalnya dengan menentukan nilai variasi konstanta *bayes* secara otomatis untuk mendapatkan akurasi yang tinggi dengan berapapun jumlah data *trainingnya*.
2. Aplikasi di sisi *client* dapat dikembangkan lebih lanjut dengan melakukan analisa terhadap *body email* bertipe html atau *mime* dan *attachment* yang dianalisa dari *contentnya*.
3. Penyaringan spam sebenarnya ada kaitannya dengan admin yang langsung menginstalasi dan mengkonfigurasi web dan mail server yang digunakan, namun bukanlah keharusan/kewajiban tetapi bila tujuannya sekedar menambah kualitas daripada pelayanannya, baguslah kiranya para administrator pada saat menginstalasi dan mengkonfigurasi software yang digunakan, mengaktifkan pengetesan bayes dan jaringan serta auto-learn dari Spam-Assassin sehingga anti spam dapat di *update* setiap saat dengan menggunakan database jaringannya tergantung kondisi yang dialami para usernya. Tetapi hal ini juga menuntut para user sendiri bisa mengerti dan memahami penggunaannya agar dapat langsung membantu admin dalam mencapai tujuannya.

DAFTAR PUSTAKA

- [1] Alimargono, 2009, Studi Banding, <http://alimargono.wordpress.com/2009/01/31/studi-banding>, diakses tanggal 27 Januari 2014.
- [2] Anugroho, P., Winarno, I., Rosyid N. M., 2009, Klasifikasi Email Spam Dengan Metode Naive Bayes Classifier Menggunakan Java Programming, Surabaya.
- [3] Fiva, R. 2009, *Langkah Mudah Administrasi Jaringan Menggunakan Linux Ubuntu 9*. Wahana Komputer, Semarang.
- [4] Graham, P. 2002. A Plan for Spam. <http://www.paulgraham.com/spam.html>.
- [5] O'Brien, C., Vogel, C., 2004, Comparing Spam-Assassin with CBDF Email Filtering, *Proceedings of the 7th Annual CLUK Research Colloquium*, Birmingham, 6-7 Januari 2004.

- [6] Pratama, A. M. R., 2008, Perancangan Dan Implementasi Mail Server Berbasis Qmail Pada Jcpanel Web Hosting Control Panel, *Seminar Nasional Aplikasi Teknologi Informasi (SNATI)*, Yogyakarta, 21 Juni 2008.
 - [7] Rachli, M., 2007, Email Filtering Menggunakan Naive Bayesian, *Tugas Akhir*, Jurusan Teknik Elektro Institut Teknologi Bandung, Bandung.
 - [8] Schwartz, A., 2004, *Spam Assassin*, O'Reilly Media, Inc., United States of America.
-