

SISTEM PEMFILTERAN GAMBAR DAN VIDEO PORNO PADA JARINGAN

Airlangga Cahya Utama, Ir. Dadet Pramadihanto, M.Eng, Ph.D

Jurusan Teknik Informatika, PENS – ITS Surabaya

Jl. Raya ITS, Surabaya

+62 (31) 594 7280; Fax: +62 (31) 594 6114

E-mail : *airlangga.cahya [at] gmail [dot] com, dadet [at] eepis-ist [dot] edu*

Abstrak

Sebagai sumber informasi, internet memberikan sumbangsih yang besar pada kehidupan modern. Sayangnya, tidak semua informasi cocok untuk semua kalangan. Sebagai contoh, informasi yang memuat perang, kekerasan, obat-obatan terlarang, dan pornografi. Beberapa hal ini tidak baik jika dikonsumsi oleh anak-anak. Berbagai macam teknik filtering digunakan untuk menyaring informasi ini. Beberapa diantaranya adalah filtering berbasis text dan nama domain. Filtering berbasis text bekerja dengan mencari kata-kata kunci dan menghitung bobotnya pada dokumen. Filtering berbasis domain menggunakan daftar domain yang dianggap berbahaya. Kedua teknik ini memiliki berbagai kelemahan. Filtering berbasis text tidak efisien jika konten yang berbahaya adalah konten gambar dan video. Sedangkan filtering berbasis domain membutuhkan sumber daya yang besar untuk melakukan pengecekan pada setiap domain yang ada. Selain itu, tingkat perubahan pada nama domain dan isinya sangatlah tinggi. Domain yang memuat konten netral dapat berubah menjadi berbahaya kapan saja. Penelitian ini menggunakan teknik pendeteksian gambar porno. Dengan pendekatan ini, filtering dilakukan pada berbagai data yang berjalan melalui jaringan. Sebuah komputer yang berdiri diantara pengguna dan internet digunakan untuk menangkap dan memproses data ini. Dengan teknik ini, tingkat pendeteksian gambar porno dapat mencapai 60%. Sedangkan tingkat false positif mencapai 10%. Proses deteksi yang digunakan pada penelitian ini memerlukan kemampuan komputansi yang tinggi. Gambar yang terdiri dari 200.000 pixel memerlukan paling sedikit 3 detik untuk memprosesnya. Penelitian lebih lanjut pada optimalisasi kecepatan diperlukan untuk mengatasi kelemahan ini.

Kata Kunci : pornografi, filter, networking, proxy, icap, haar, eigenface, image

1. Pendahuluan

Dengan berkembangnya teknologi komunikasi di negara kita, kini semakin mudah dan murah bagi masyarakat untuk menikmati internet. Internet memberikan pengetahuan yang beragam bagi penggunanya. Sayangnya tidak semua materi termasuk dalam kategori aman untuk dikonsumsi kalangan umum. Salah satunya adalah materi pornografi. Materi pornografi beredar cepat di internet, dan untuk mencarinya tidak perlu susah payah dan tidak perlu mengeluarkan banyak uang.

Untuk mencegahnya diperlukan suatu sistem filtering materi yang berasal dari internet. Metode yang sering dipakai saat ini adalah filtering berdasarkan alamat domain atau berdasarkan tulisan yang ada dalam website tersebut. Namun pendekatan ini tidak akurat karena internet bersifat dinamis. Perubahan terjadi di internet bahkan dalam hitungan detik. Hal ini tentu merepotkan untuk merawat dan memperbaharui daftar panjang domain yang dicurigai. Kelemahan ini mendasari munculnya konsep untuk menyeleksi situs berdasarkan gambar dan video yang dimuat. Metode ini meningkatkan akurasi dalam menyaring konten pornografi dan mengurangi kerja system administrator.

Data yang diterima dari internet terlebih dahulu melalui beberapa proses penyaringan sebelum dilanjutkan kepada user. Pertama-tama, data yang diterima diseleksi

terlebih dahulu apakah data tersebut memuat informasi gambar atau tidak. Jika data tersebut berupa gambar, maka akan dilewatkan pada proses pendeteksian gambar pornografi. Proses deteksi gambar porno terdiri dari proses skin detection yaitu proses pengelompokan warna kulit, texture analysis yaitu proses analisa tekstur kulit manusia, object identification yaitu proses identifikasi bagian-bagian tubuh manusia, image classification yaitu proses menentukan apakah gambar tersebut termasuk ke dalam materi pornografi atau bukan, kemudian tahap terakhir yaitu melakukan blocking pada data apabila memenuhi kriteria sebagai gambar pornografi.

Penelitian ini melanjutkan penelitian sebelumnya yang telah mencapai tingkat pendeteksian. Dengan mengintegrasikan fungsi pendeteksian gambar pada jaringan, khususnya proxy server, maka program yang telah dibuat akan menjadi lebih berguna dan bermanfaat.

2. Teori Penunjang

2.1. Jaringan

pktd: A Packet Capture and Injection Daemon yang bertindak sebagai daemon pada jaringan. pktd menangkap paket yang lewat pada kernel, kemudian mengolahnnya apakah memenuhi kriteria tertentu, dan kemudian memanipulasinya. pktd diimplementasikan dalam 12.000 baris kode C diatas pcap dan libnet. pktd telah di porting

dan dapat dijalankan pada sistem operasi Linux, Solaris, dan FreeBSD.

pktd telah dioptimalkan untuk bekerja pada lingkungan dengan kebutuhan transfer data tinggi hingga mencapai 832 Mbps (159MB dalam 1,53 detik) yang terdiri dari 113.120 paket UDP. Dalam penelitian ini tool tcpdump yang berjalan pada sistem FreeBSD tidak mampu menangkap seluruh header paket yang lewat dan kehilangan sekitar 0,2 % data.

Penelitian lain seperti nCap: Wire-speed Packet Capture and Transmission memberikan kemungkinan untuk mengakses paket secara mentah (raw) pada memori kernel. Dengan ukuran paket 64 byte, nCap mampu memproses kurang lebih 205 Mbps atau sekitar 561.078 paket per detik.

2.2. Warna Kulit

Rigan Ap-apid membuat sebuah sistem yang menggunakan beberapa macam color space, diantaranya yaitu RGB, Normalized RGB, dan HSV sekaligus untuk menentukan apakah warna tersebut termasuk kedalam warna kulit atau bukan. Hasilnya adalah pencapaian pendeteksian hingga 96% dengan false positives 8%.

An Objectionable Image Detection System based on Region of Interest mengemukakan cara yang unik menggunakan Region of Interest untuk menemukan area kulit. Gambar terlebih dahulu dilewatkan pada algoritma gaussian untuk memperhalus gambar, setelah itu dicari intensitasnya, komposisi merah-hijau, komposisi biru-kuning, dan kontras warna kulit. Setelah itu dibuat saliency mapnya untuk selanjutnya diproses menggunakan ROI Extraction.

Sistem Seleksi Gambar Khusus Dewasa Melalui Web Proxy Menggunakan Image Processing menggunakan colos space RGB dan untuk metode filter warna kulit menggunakan model Peer yang telah dimodifikasi. Berikut hasil modifikasi yang diterapkan dalam bentuk pseudo code $r > 95$ and $g > 40$ and $b > 20$ and $\max(r, g, b) - \min(r, g, b) > 15$ and $\text{abs}(r - g) > 15$ and $r > g$ and $r > b$.

2.3. Deteksi obyek

Deteksi Pornografi Pada Citra Digital Menggunakan Pengolahan Citra dan Jaringan Syaraf Tiruan menggunakan deteksi tepi Sobel. Hasil dari deteksi tepi Sobel ini kemudian dicari polanya menggunakan pembelajaran Learning Vector Quantization.

Sistem Seleksi Gambar Khusus Dewasa Melalui Web Proxy Menggunakan Image Processing menggunakan fitur Haar-like yang terdapat pada library OpenCV. Setiap fitur Haar terdiri dari gabungan kotak hitam dan kotak putih dalam berbagai macam variasi posisi.

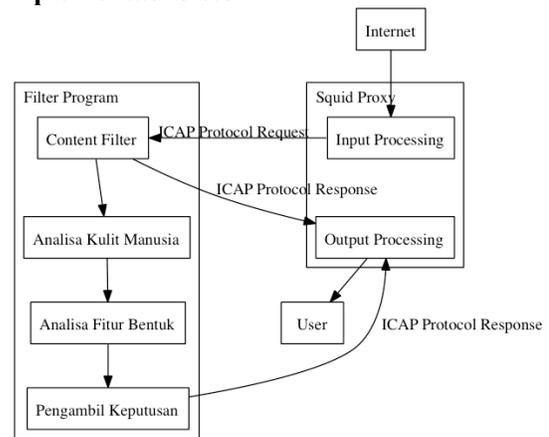
2.4. Pornografi

Beberapa definisi dari kata pornografi diatas memiliki sebuah kesamaan makna, yaitu materi erotis yang dibuat untuk membangkitkan nafsu berahi. Makna ini rancu dan sulit diterapkan karena masing-masing orang memiliki

pandangan yang berbeda-beda terhadap erotis. Apalagi sampai pada “membangkitkan” yang tentu saja berbeda-beda pada setiap orang. Oleh karena itu diambil sebuah kesimpulan sederhana dari pornografi :

- Materi pornografi adalah gambar yang memuat obyek manusia telanjang hingga terlihat alat kelamin tanpa terhalang oleh benda.

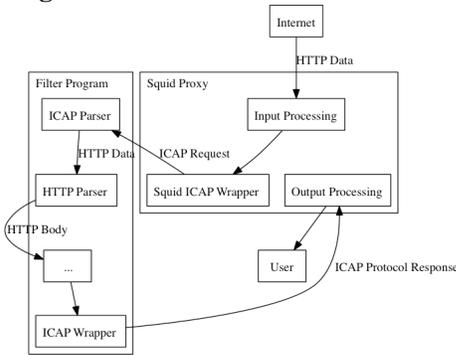
3. Implementasi Sistem



Gambar 3.1 System Overview

Secara garis besar, sistem terbagi menjadi content filter, analisa kulit manusia, analisa fitur bentuk, dan pengambil keputusan. Dimana content filter akan menangani data HTTP dan ICAP untuk dideteksi tipe data yang berada didalamnya. Analisa kulit manusia akan menfilter bagian kulit manusia dengan gambar latar belakang yang mengganggu proses deteksi. Setelah itu memasuki proses analisa fitur bentuk yang akan memproses dan mencari obyek yang mengandung unsur pornografi. Setelah itu akan diambil keputusan dan hasil dari filtering akan diteruskan kepada proxy server untuk dilanjutkan kepada pengguna. Pada bagian Content Filter akan melibatkan protokol ICAP, pembacaan header file. Pada bagian analisa kulit manusia akan ada filtering warna menggunakan model Peer yang dimodifikasi. Pada analisa fitur bentuk akan menggunakan fitur Haar-like yang digabung dengan Ada Boosting dan Cascade Classifier. Sistem pengambil keputusan menggunakan sistem penambahan sederhana, dimana jika sistem mendeteksi adanya obyek maka jumlah obyek akan bertambah.

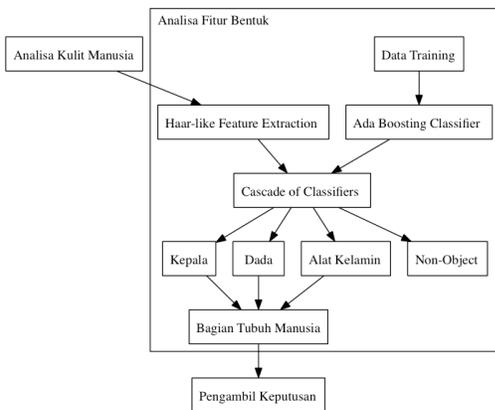
3.1. Jaringan



Filtering pada jaringan dilakukan dengan memanfaatkan Squid proxy server versi 3 yang akan di konfigurasi untuk bekerja sebagai ICAP client dan mengalihkan semua percakapan HTTP yang dilakukan antara user dan web-server.

Komunikasi dilakukan menggunakan protokol ICAP.

3.2. Deteksi Gambar



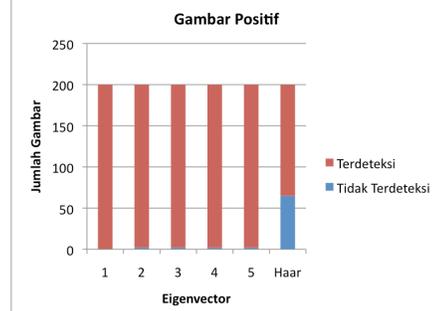
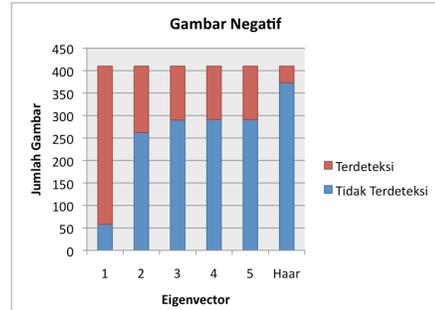
Algoritma yang dipakai adalah Haar like feature. Algoritma haar-wavelet adalah algoritma yang pada dasarnya adalah melakukan operasi penjumlahan pixel pada satu area tertentu, untuk kemudian dikurangi dengan jumlah pixel pada satu area lain. Dari berbagai macam operasi haar-wavelet dapat didapatkan berbagai macam fitur dari gambar yang diambil.

4. Hasil dan Analisa

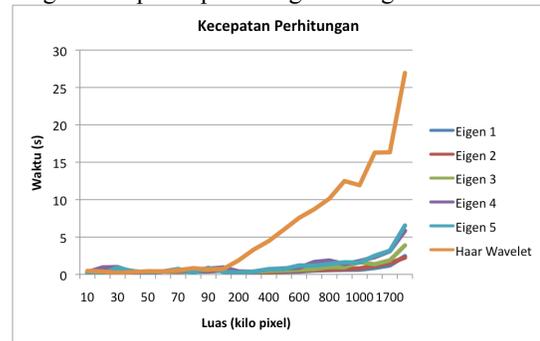
4.1. Deteksi Gambar

Pada test digunakan 410 gambar bukan gambar pornografi yang terdiri dari berbagai macam ukuran, warna, dan obyek. Untuk gambar pornografi digunakan 200 gambar sample berbagai ukuran. Gambar yang digunakan lebih besar dari 64x64 pixel.

Hasil pendeteksian menunjukkan hasil sebagai berikut :

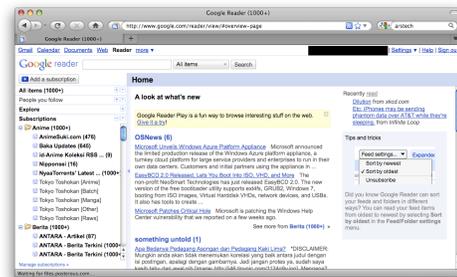


Dengan kecepatan perhitungan sebagai berikut :



4.2. Website

Efek yang didapat pada saat membuka website melalui program filter ini tidaklah terlihat, selain dari penambahan waktu tunggu. Beberapa website dengan penggunaan kode javascript yang banyak tetap dapat di load dengan benar dan baik. Situs YouTube masih dapat ditampilkan walaupun terdapat pengurangan kemampuan streaming hingga 0%. Sehingga video tidaklah distreaming, tetapi dikirim kepada client secara penuh setelah tertahan pada program filter.





5. Kesimpulan

Dari hasil diatas didapatkan proses pendeteksian masih belum cukup akurat dan lambat. Saran yang dapat diajukan untuk mengatasi masalah ini adalah dengan menggunakan algoritma pendeteksian lain. Untuk permasalahan kecepatan, peningkatan dapat dilakukan dengan mengubah kode agar dapat berjalan pada sistem GPU atau sistem cluster.

Untuk kemampuan menampilkan website dapat diambil kesimpulan tidak ada website yang terganggu kerjanya selama website tersebut tidak memanfaatkan teknologi streaming. Untuk mengatasi masalah ini haruslah dilakukan perubahan pada desain dasar aplikasi ini, dikarenakan teknik buffering yang dipakai, sehingga semua data di tampung terlebih dahulu oleh server.

6. Daftar Pustaka

- [1] Alberto Albiol, Luis Torres, Edward J. Delp, Optimum Color Spaces for Skin Detection, 2001
- [2] J. Elson, A.Cerpa, RFC 3507 : Internet Content Adaptation Protocol (ICAP), 2003
- [3] Jose Maria Gonzalez, Vern Paxson, pktd: A Packet Capture and Injection Daemon, 2004
- [4] Luca Deri, nCap: Wire-speed Packet Capture and Transmission, 2005
- [5] Margaret M. Fleck, David A. Forsyth, Chris Bregler, Finding Naked People, 1996
- [6] Muhammad Fatoni Ashariyanto, Rudy Cahyadi Hario Pribadi, Dadet Pramadihanto, Sistem Seleksi Gambar Khusus Dewasa Melalui Web Proxy Menggunakan Image Processing, 2007
- [7] Nazrul Effendy, Rifqi Imanto, Ayodya P.T, Deteksi Pornografi Pada Citra Digital Menggunakan Pengolahan Citra dan Jaringan Syaraf Tiruan, 2008
- [8] Rainer Lienhart, Alexander Kuranov, Vadim Pisarevsky, Empirical Analysis of Detection Cascades of Boosted Classifiers for Rapid Object Detection, 2002
- [9] Wei Zeng, Wen Gao, Tao Zhang, Yang Liu, Image Guarder: An Intelligent Detector For Adult Images, 2003

[CV Penulis]

Airlangga Cahya Utama, menyelesaikan kuliah D4 Teknik Informatika Politeknik Elektronika Negeri Surabaya-ITS (PENS-ITS).