

OPTIMALISASI WEB PROXY CACHE UNTUK MENINGKATKAN KECEPATAN AKSES INTERNET MENGGUNAKAN ID3

Muhammad Riza Hilmi¹, I Made Darma Susila²

^{1,2}Program Studi Sistem Komputer Institut Teknologi dan Bisnis STIKOM Bali
Email: riza@stikom-bali.ac.id¹

ABSTRAK

Pertukaran data di Internet semakin terus meningkat secara signifikan, terutama penyediaan informasi yang diakses melalui website. Hal ini menyebabkan tingginya trafik resource pada server web dan memerlukan bandwidth yang besar untuk koneksi pada lingkungan host pengakses sehingga jika bandwidth terbatas maka host pengakses akan mengalami akses Internet yang lambat. Penyediaan proxy yang memiliki fasilitas cache sering digunakan untuk mengurangi beban tersebut. Akan tetapi proxy cache masih memiliki beberapa kelemahan diantara hanya sebagian kecil resource web yang disimpan dalam cache dan yang tersimpan di cache sebagian besar kurang dapat dimanfaatkan sehingga jika pengguna mengakses kembali website masih memerlukan bandwidth yang besar. Penelitian ini menggunakan metode Iterative Dychotomizer version 3 (ID3) untuk membantu mengoptimalkan penyimpanan resource dalam cache. ID3 memberikan informasi tentang objek yang akan disimpan ke dalam cache yang dianggap penting sebagai input dan output jaringan digunakan untuk memandu pemilihan objek cache yang diakses oleh pengguna. Hasil penelitian ini menunjukkan bahwa ID3 dapat membantu mengoptimalkan kemampuan web proxy guna menyimpan data yang sering diakses pengguna kedalam cache dan mengoptimalkan penggunaan bandwidth sehingga dapat mempercepat akses internet oleh pengguna.

Kata kunci: web proxy, cache proxy, optimalisasi cache.

ABSTRACT

Data exchange on the Internet continues to increase significantly, especially the provision of information accessed through the website. This causes high resource traffic on web servers and requires large bandwidth to connect to the access host environment so that if bandwidth is limited then accessing hosts will slow Internet access. Provision of proxies that have cache facilities is often used to reduce the burden. However, the proxy cache still has some disadvantages between only a small portion of web resources stored in the cache and most of the cache stored can not be utilized so that if the user accesses the website again, it still requires large bandwidth. This study uses the Iterative Dychotomizer version 3 (ID3) method to help optimize storage of resources in the cache. ID3 provides information about objects to be stored in cache which are considered important as input and output are used to guide the selection of cache objects accessed by users. The results of this study indicate that ID3 can help optimize the ability of web proxies to store data that is frequently accessed by users into the cache and optimize bandwidth usage so that users can speed up internet access.

Keywords: web proxy, cache proxy, cache optimization

1. Pendahuluan

Kemajuan dibidang teknologi informasi semakin pesat, terutama dalam hal pertukaran data menggunakan Internet. Internet merupakan fasilitas yang sering digunakan untuk berbagai kebutuhan diantaranya sebagai media pembelajaran, berbagi informasi, media komunikasi dan bahkan bermain *game*. Oleh karena itu penggunaan internet yang stabil sangat diperlukan untuk pengguna agar internet dapat diakses secara optimal dan lancar.

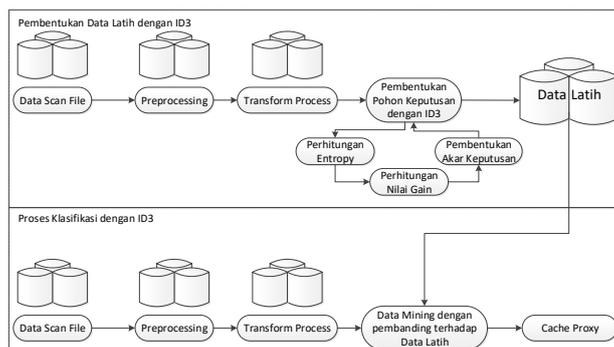
Berbagai macam cara dilakukan untuk meningkatkan kelancaran mengakses internet, diantaranya meningkatkan *bandwidth*, membangun *router* dan *server-server* pendukung lainnya, salah satunya adalah membangun *server proxy*. *Server proxy* dirancang dengan tiga tujuan yaitu mengurangi *bandwidth*, mengurangi keterlambatan yang dirasakan pengguna, dan mengurangi beban pada *server* asal dengan menyalin salinan objek *website* (Imtiaz, 2014). Untuk mencapai tujuan ini, *proxy server* menggunakan teknik *cache*. *Caching web* adalah mekanisme yang digunakan untuk meningkatkan kinerja jaringan dengan mengurangi lalu lintas data pada jaringan (Sathiyamoorthi, 2012), memuat objek di *server web* dan menunda mengakses halaman *web* dengan menyimpan halaman *web* yang sering diakses pada *cache proxy* yang ditempatkan dalam jaringan (Niranjan, 2013).

Namun pada umumnya *proxy cache* masih memiliki beberapa kelemahan diantara hanya sebagian kecil *resource web* yang disimpan dalam *cache* (Zhaolei, 2009) dan yang tersimpan di *cache* sebagian besar kurang dapat dimanfaatkan sehingga jika pengguna mengakses kembali *website* masih memerlukan *bandwidth* yang besar (Ponnusamy, 2012).

Oleh karena itu, penulis melalui penelitian ini mengusulkan teknik *caching* yang lebih baik dimana akan terjadi pemilihan data yang perlu disimpan ataupun yang tidak perlu disimpan pada *proxy server* dengan metode *Iterative Dychotomizer version 3 (ID3)*. *ID3* digunakan sebagai metode untuk menyelesaikan masalah klasifikasi data yang akan disimpan dalam *cache proxy*. Metode *ID3* dapat mengubah pengambilan keputusan yang awalnya kompleks dan global menjadi lebih sederhana dan spesifik (Suyanto, 2014). Penerapan *ID3* diharapkan mampu membantu dalam proses klasifikasi objek *file web* untuk pemilihan *file web* atau data yang akan disimpan dalam *cache proxy*. Dengan menggunakan *ID3* untuk membantu mengoptimalkan penyimpanan *resource* dalam *cache* diharapkan dapat mengoptimalkan kemampuan *web proxy* guna menyimpan data yang sering diakses pengguna kedalam *cache* dan mengoptimalkan penggunaan *bandwidth* sehingga dapat mempercepat akses internet oleh pengguna.

2. Metode

Pada penelitian ini menggunakan teknik penyimpanan pada *cache proxy* yang lebih baik dimana akan terjadi proses klasifikasi *file website* atau data yang perlu disimpan dalam *proxy* menggunakan metode *Iterative Dychotomizer version 3 (ID3)*. Metode *ID3* digunakan sebagai metode untuk menyelesaikan masalah klasifikasi *file website* atau data yang pada mulanya rumit dan bersifat umum menjadi lebih sederhana dan spesifik. Gambaran perancangan sistem penyimpanan pada *cache proxy* menggunakan *ID3* secara umum dapat dilihat dari gambar 1.



Gambar 1. Sistem penyimpanan *cache proxy* dengan klasifikasi *ID3*

Pada gambar 1 menunjukkan bahwa pada pembentukan data latih, sistem akan mendapatkan *scan data* dari *file website* yang diakses oleh pengguna kemudian akan dilakukan proses *preprocessing*, *transform process* hingga pembentukan pohon keputusan dengan *ID3* sehingga mendapatkan data latih yang optimal. Semakin sering pengguna mengakses website yang sering dikunjungi akan semakin optimal data yang akan didapatkan dan disimpan pada *cache proxy*.

3. Hasil dan Pembahasan

a. Pembentukan Data Latih

Pembentukan data latih pada proses akuisisi data dilakukan dengan melakukan *scan data* menggunakan aplikasi *squid3* pada *server proxy* dengan mengakses 5 website yang telah ditentukan untuk mendapatkan *data* yang *valid* sebagai data latih, yaitu *data* yang memiliki kriteria terhadap masing-masing atribut penilaian. *File* yang didapat dari proses *scan data* di-export ke dalam bentuk *file log*. Pada proses *preprocessing* menghasilkan 18.502 jumlah *data* yang bersifat unik (memiliki perbedaan isi dari atribut yang telah ditentukan dalam sistem klasifikasi) dengan waktu yang diperlukan adalah 24 menit 12 detik. *Transform proses* melakukan penandaan dari *record data* yang sudah dibentuk, penandaan dengan pemberian *class label* yang dimaksud adalah *class label file website* yang memiliki *JobId* untuk masing-masing *website* (yang diidentifikasi dengan *ClientId*). Proses penandaan bertujuan untuk membentuk ciri dari masing-masing data pada *record data* latih. Pada proses penghitungan *ID3*, hasil penghitungan dari nilai *entropy* dan *gain* dari masing-masing atribut tersebut menunjukkan nilai *gain* yang paling besar adalah pada nilai atribut *Modification*, sehingga atribut *Modification* adalah sebagai akar *node* keputusan jenis *file website* yang di simpan dalam *cache*. Hasil akhir pembentukan *node* keputusan yang dibentuk seperti gambar 2.

```

cache/
├── id3
│   └── modification
│       ├── accessibility
│       │   ├── fileid3.log
│       │   └── readable
│       │       └── fileid3.log
│       └── fileid3.log
└── 4 directories, 3 files
    
```

Gambar 2. Hasil Pohon Keputusan *File Website*

Pohon keputusan yang dihasilkan dari penghitungan *ID3* untuk penyimpanan *file website* pada *cache proxy* menunjukkan pembentukan *node* terbanyak adalah sebanyak 3 *node*. *Node* yang mencapai kedalaman 3 *node* adalah menunjukkan keputusan *modification*, *accessibility* dan *readable*.

b. Klasifikasi ID3

Data uji yang diujikan adalah data uji yang diambil dari beberapa objek studi kasus. Seperti data uji *website youtube.com*, *stikom-bali.ac.id*, *elearning.stikom-bali.ac.id*, *yahoo.com* dan *facebook.com*. Hasil dari pengambilan data uji dan *load* ke dalam *cache proxy* dapat ditunjukkan pada tabel 1.

Tabel 1. Hasil Pengambilan Data Uji *Cache Proxy* tanpa ID3

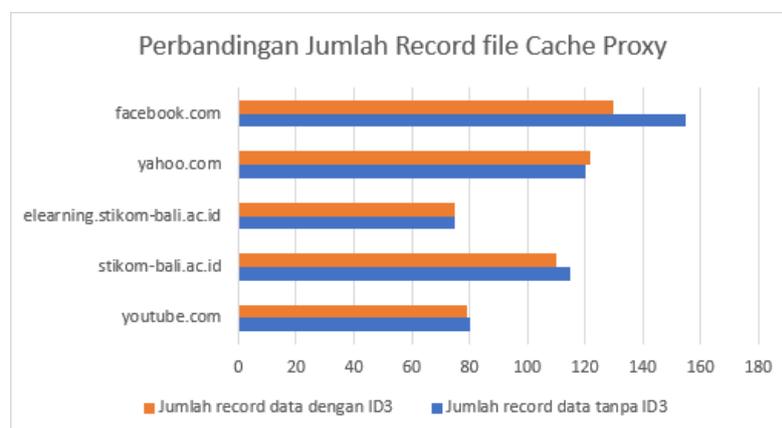
No	Nama Website	Ukuran log file (bytes)	Waktu pengambilan file (second)	Jumlah record data
1	<i>youtube.com</i>	710	80	80
2	<i>stikom-bali.ac.id</i>	632	60	115
3	<i>elearning.stikom-bali.ac.id</i>	658	75	75
4	<i>yahoo.com</i>	1260	105	120
5	<i>facebook.com</i>	1691	143	155
Jumlah Data total			463	545

Tahap *preprocessing* melakukan filterisasi terhadap *record data*. Pada tahap *transform process* data akan dikelompokkan dengan kesamaan sejumlah atribut yang dimiliki yaitu dari atribut *Readable*, *Accessibility*, dan *Modification* dengan tujuan untuk mencari nilai pola masukan (P) dan target (T) yang unik dari sejumlah data *file scan*. Pengujian klasifikasi ID3 dilakukan dengan menggunakan data uji seperti ditunjukkan pada tabel 2.

Tabel 2. Hasil Pengambilan Data Uji *Cache Proxy* dengan ID3

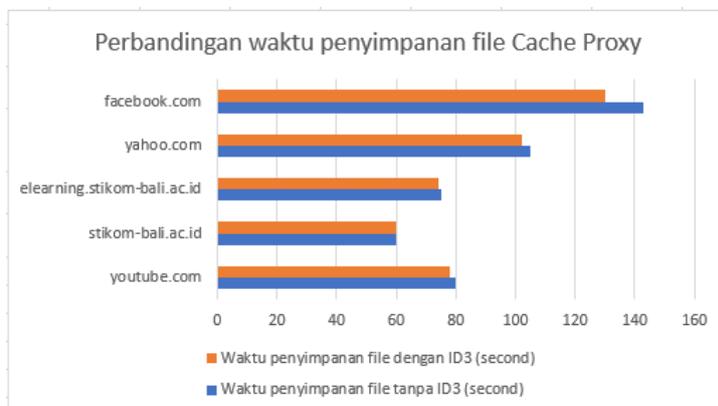
No	Nama Website	Ukuran log file (bytes)	Waktu pengambilan file (second)	Jumlah record data
1	<i>youtube.com</i>	680	78	79
2	<i>stikom-bali.ac.id</i>	625	60	110
3	<i>elearning.stikom-bali.ac.id</i>	634	74	75
4	<i>yahoo.com</i>	960	102	122
5	<i>facebook.com</i>	1391	130	130
Jumlah Data total			444	516

Klasifikasi yang dilakukan terhadap sampel data uji pada tabel 2, yaitu mengklasifikasikan terhadap data uji sejumlah 516 data *record*, sehingga dapat digambarkan perbandingannya seperti pada gambar 3.



Gambar 3. Perbandingan Jumlah Record File Website pada *Cache Proxy*

Pengujian jaringan menunjukkan bahwa *ID3* untuk klasifikasi data yang akan disimpan pada *cache proxy* adalah sebesar 94,68% dibandingkan dari tanpa klasifikasi dengan *ID3*. Sedangkan untuk perbedaan waktu penyimpanan *file website* pada *cache proxy* dapat ditunjukkan pada gambar 4.



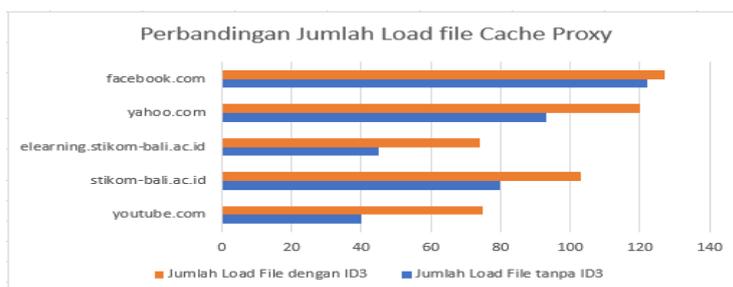
Gambar 4. Perbandingan Waktu Penyimpanan *File Website* pada *Cache Proxy*

Pada gambar 4 menunjukkan bahwa waktu yang dibutuhkan untuk menyimpan *file website* pada *cache proxy* mengalami percepatan waktu sebesar 19 detik. Pengujian berikutnya adalah perbandingan *load file* yang tersimpan didalam *cache proxy*. Semakin banyak jumlah *file* yang digunakan untuk menampilkan halaman *web* maka semakin optimal *cache proxy*, perbandingannya seperti ditunjukkan pada tabel 3.

Tabel 3. Hasil Perbandingan *Load File Cache Proxy* dengan *ID3* dan tanpa *ID3*

No	Nama Website	Jumlah Load File tanpa ID3	Jumlah Load File dengan ID3
1	youtube.com	40	75
2	stikom-bali.ac.id	80	103
3	elearning.stikom-bali.ac.id	45	74
4	yahoo.com	93	120
5	facebook.com	122	127
Jumlah Data total		380	499

Dari tabel 3 dapat diketahui bahwa jumlah *load file* yang paling banyak adalah *cache proxy* yang menggunakan *ID3*, hal ini menunjukkan bahwa *cache proxy* menggunakan *ID3* lebih optimal dalam menyimpan dan menggunakan *file website*. Jika dilihat dalam bentuk grafik ditunjukkan pada gambar 5.



Gambar 5. Perbandingan *Load File Website* pada *Cache Proxy*

Pengujian selanjutnya yaitu kecepatan akses pengguna terhadap website yang telah ditentukan dalam data uji menggunakan klasifikasi *ID3* dibandingkan tanpa klasifikasi *ID3* dengan hasil seperti pada tabel 4.

Tabel 4. Hasil Perbandingan Kecepatan Akses Website

No	Nama Website	Waktu akses tanpa <i>cache proxy</i>	Waktu akses dengan <i>cache proxy</i> tanpa <i>ID3(second)</i>	Waktu akses dengan <i>cache proxy ID3(second)</i>
1	<i>youtube.com</i>	28	20	10
2	<i>stikom-bali.ac.id</i>	29	25	15
3	<i>elearning.stikom-bali.ac.id</i>	12	7	5
4	<i>yahoo.com</i>	33	25	22
5	<i>facebook.com</i>	35	20	14
Rata-rata waktu akses		27,4	19,4	13,2

Pada tabel 4 dapat dilihat bahwa rata-rata waktu akses website yang paling cepat adalah yang menggunakan *server proxy* dengan optimalisasi *ID3* pada *cache proxy* yaitu sebesar 13,2 detik, sehingga *ID3* dapat mempercepat akses internet dibandingkan *cache proxy* tanpa *ID3* atau bahkan tanpa menggunakan *cache proxy*. Jika digambarkan dalam bentuk grafik dapat ditunjukkan pada gambar 6.



Gambar 6. Perbandingan Waktu Akses

4. Simpulan

Dari pengujian sistem yang telah dilakukan, dapat disimpulkan bahwa *cache proxy* yang menggunakan *ID3* dapat mempercepat akses internet dengan hasil pengujian rata-rata waktu akses website sebesar 13,2 detik dibandingkan *cache proxy* tanpa *ID3* atau bahkan tanpa menggunakan *cache proxy*. Selain itu, jumlah *load file* yang paling banyak adalah *cache proxy* yang menggunakan *ID3*, hal ini menunjukkan bahwa *cache proxy* menggunakan *ID3* lebih optimal dalam menyimpan dan menggunakan *file website* dibandingkan dengan *cache proxy* tanpa *ID3*. Sedangkan dalam hal waktu penyimpanan *file* ke dalam *cache proxy* lebih cepat menggunakan *ID3* dibandingkan tanpa menggunakan *ID3*. Dan jumlah *record data* pada *ID3* untuk klasifikasi data yang akan disimpan pada *cache proxy* adalah sebesar 94,68% dibandingkan dari tanpa *ID3*.

Daftar Rujukan

- Intiaz, A. and Hossain, M. J. 2014. "Distributed cache management architecture: To reduce the internet traffic by integrating browser and proxy caches". International Conference on Electrical Engineering and Information & Communication Technology, Dhaka, pp. 1-4.
- Niranjan, Y. Tiwari, S. and Gupta, R. 2013. "Average memory access time reduction in multilevel cache of proxy server", 3rd IEEE International Advance Computing Conference (IACC), Ghaziabad, pp. 44-47.
- Ponnusamy, S. P. and Kathikeyan, E. 2012. "Cache optimizatio Hot-Point Proxy (HPPProxy) using Dual Cache Replacement Policy". International Conference on Communication and Signal Processing, Chennai, pp. 108-113.
- Sathiyamoorthi, V. and Murali Bhaskaran, V. 2012. "Web caching through modified cache replacement algorithm". International Conference on Recent Trends in Information Technology, pp. 483-487. doi: 10.1109/ICRTIT.2012.6206749.
- Suyanto. 2014. "Artificial Intelligence (Seraching, Reasoning, Planning, Learning) Revisi Kedua". Bandung: Informatika.
- Zhaolei Duan, Zhimin Gu and Xiaoguang Ding, 2009. "WPCC: A novel web proxy cache cluster". International Conference on Advanced Communication Technology, Phoenix Park, pp. 2205-2208.