

SISTEM PENAPISAN UNTUK MENGATASI ANCAMAN INTERNET DALAM PENDIDIKAN

Oleh

HAMIRUL'AINI BT. HAMBALI
HAZARUDDIN BIN HARUN

Sekolah Teknologi Maklumat
Universiti Utara Malaysia
06010 Sintok, Kedah.
Tel : 04-9284605

ABSTRAK

Internet merupakan satu media yang popular pada masa ini kerana banyak membantu pengguna untuk mencari dan menyebarkan maklumat. Walau bagaimanapun, terdapat maklumat dalam Internet yang tidak sesuai dan berbahaya terutama bagi golongan pelajar. Bagi menghalang pelajar daripada mencapai bahan-bahan negatif ini, industri perisian telah mengambil langkah membangunkan sistem perisian yang boleh menapis kandungan Internet. Bagaimanapun perisian penapisan yang sedia ada masih tidak dapat menjalankan fungsinya dengan baik di mana ia gagal menapis seratus peratus bahan yang tidak sepatutnya tetapi pada masa yang sama ia menahan beberapa peratus bahan yang sepatutnya boleh diterima. Perisian yang ada juga lebih menumpukan kepada penapisan bahan yang ditulis dalam Bahasa Inggeris sedangkan bahan yang tidak sesuai untuk pelajar terdapat juga dalam Bahasa Melayu. Oleh itu, kajian ini telah membangunkan perisian penapisan Internet yang menggunakan pendekatan berasaskan agen dan mekanisme gabungan beberapa katakunci Bahasa Melayu bagi membolehkan ia berfungsi dengan lebih berkesan. Katakunci yang digunakan dalam menentukan halaman mana yang perlu ditapis atau dilepaskan telah dipecahkan kepada dua kategori iaitu katakunci 'halang' dan katakunci 'lepas'. Perisian dalam kajian ini mengaplikasikan ciri-ciri agen dan seterusnya dapat menapis atau melepaskan kandungan Internet yang mengandungi katakunci-katakunci tersebut.

PENGENALAN

Semenjak Internet mula diperkenalkan kepada umum, ia telah menjadi satu media yang popular kerana membolehkan manusia berhubung, memperolehi dan menyalurkan pelbagai maklumat dengan pantas tanpa mengira agama dan ideologi politik sesebuah negara. Internet yang didefinisikan sebagai rangkaian komputer besar yang merangkumi perhubungan seluruh dunia telah banyak membantu manusia di dalam pelbagai bidang terutama sekali bidang pendidikan. Ia menyediakan pelbagai faedah dan kemudahan kepada pengguna untuk mencari sumber maklumat, berkomunikasi sesama manusia dan menyebarkan maklumat.

Internet mempunyai kelebihan berbanding dengan media penyiaran yang lain seperti televisyen dan radio kerana pengguna tidak perlu melalui badan tapisan seperti yang

dilakukan oleh badan penyiaran antarabangsa atau negara tertentu. Pengguna akan dapat menikmati segala maklumat yang ada kerana jika melalui badan penapisan, kemungkinan maklumat tersebut hanya terhad mengikut kehendak dan ideologi sesuatu organisasi atau negara.

Namun demikian, maklumat yang ada di Internet tidak semuanya berfaedah dan baik untuk pengguna. Internet juga turut menyebarkan bahan lucah, kegiatan subversif pengganas, ideologi songsang dan bahan-bahan negatif lain yang tidak sesuai dan merbahaya terutama bagi capaian golongan pelajar atau remaja bawah umur.

Pada tahun 1994, kebimbangan terhadap bahan-bahan tidak sesuai untuk pelajar dan remaja yang mudah didapati di Internet telah mendapat perhatian umum terutama di Amerika (Turow, 1999). Dalam satu bancian oleh USA Today/CNN pada bulan Mei 1999, didapati 65% remaja mengatakan internet merupakan antara penyumbang kepada terjadinya keganasan seperti yang berlaku di Littleton, Colorado. Dalam kejadian itu dua pelajar remaja berusia 18 dan 17 tahun telah menembak mati 13 orang pelajar dan seorang guru di sekolah mereka. Menurut Bureau of Alcohol, Tobacco and Firearms, agen pusat Amerika Syarikat telah mengenalpasti sekurang-kurangnya 30 kejadian bom dan 4 cubaan mengebom antara 1985 dan Jun 1996 berlaku kerana kebanyakan yang disyaki telah mendapatkan literasi membuat bom dari Internet. Dalam bulan Februari 1996, tiga pelajar sekolah tinggi di Syracuse, New York telah dituduh atas kesalahan membuat bom buatan sendiri berdasarkan plan yang mereka dapat dari Internet.

Tidak berapa lama selepas kejadian di Colorado, Timbalan Presiden Amerika Syarikat, Al Gore menyatakan bahawa apabila Internet digunakan, ibu bapa perlu menggunakan teknologi penapisan untuk menghadkan capaian ke halaman-halaman yang tidak baik (Gore, 1999). Hasil kaji selidik yang dibuat oleh NetValue pada bulan Mac 2001 pula melaporkan seorang dari lima kanak-kanak di Britain mengunjungi laman-laman lucah dengan purata masa yang dihabiskan melawat laman tersebut adalah 28 minit.

KEPERLUAN SISTEM PENAPISAN INTERNET

Berikutan situasi di atas, industri perisian telah mengambil langkah untuk mengurangkan masalah ini dengan membangunkan sistem perisian yang boleh menapis bahan-bahan daripada internet. Sistem penapisan atau *filtering system* merupakan satu sistem yang dapat mengelakkan atau menghalang pengguna daripada mencapai atau memuat turun (*download*) bahan-bahan yang tidak sesuai. Apabila sesuatu sistem penapisan ini dihidupkan, pengguna tidak dapat memasuki laman-laman yang tidak sesuai atau *illegal sites*.

Clinton ada menyatakan bahawa teknologi penapisan dan sekatan internet yang canggih adalah kaedah yang lebih efektif dalam menghalang pelajar daripada mencapai bahan yang tidak sesuai berbanding dengan sebarang penggubalan undang-undang (Clinton, 1997). Tindakan Kongres Undang-Undang Amerika Syarikat mengarahkan sekolah-sekolah dan perpustakaan menggunakan perisian penapisan seharusnya dijadikan contoh oleh negara-negara lain. Ini adalah kerana golongan remaja dan kanak-kanak yang bersekolah tidak boleh didedahkan kepada perkara-perkara negatif. Mereka selalunya cenderung untuk mengikut atau meniru apa yang mereka lihat. Dengan menggunakan penapis Internet, peratusan mereka untuk mencapai halaman web yang berunsur negatif amatlah rendah.

Di Malaysia pula, Menteri Tenaga, Telekomunikasi dan Multimedia, Datuk Amar Leo Moggie telah mengajak rakyat Malaysia ke arah kesedaran bagi menggunakan Internet untuk tujuan baik. Beliau juga telah mencadangkan supaya sistem penapisan Internet digunakan secara meluas termasuk dalam bidang pendidikan dan juga korporat (Utusan Malaysia, 14/4/99). Penggunaan sistem ini bertujuan untuk mewujudkan satu persekitaran penggunaan sumber Internet yang bersih dari segala unsure negatif.

MASALAH SISTEM PENAPISAN SEDIA ADA

Daripada laporan Consumer Reports, perisian-perisian yang dikatakan dapat menapis Internet untuk menghalang kanak-kanak dan pelajar daripada bahan yang tidak dibenarkan dan lucu masih tidak dapat menjalankan fungsinya dengan baik (Consumer Report, March 2001). Terdapat dua masalah utama dengan perisian penapisan yang ada di pasaran di mana pertama; ada sesetengah perisian penapisan yang tidak menapis dengan sempurna, kedua; ada perisian penapisan yang terlalu sempurna sehingga tertapis halaman yang berfaedah.

Berdasarkan kajiannya ke atas 6 perisian penapisan iaitu Cyber Patrol, Cybersitter 2000, Cyber Snoop, Internet Guard Dog, Net Nanny dan Norton Internet Security 2001, didapati Cybersitter 2000 dan Internet Guard Block Dog boleh menahan satu daripada lima halaman yang tidak baik. Net Nanny dan Cyber Snoop pula masing-masing hanya dapat menapis sebanyak 48 % dan 10 % halaman yang tidak sepatutnya dilepaskan. Perisian yang paling mahal iaitu Norton Internet Security dapat menapis sebanyak 80 % halaman tetapi pada masa yang sama menahan halaman kerajaan berkaitan dengan penyalahgunaan dadah.

Adalah didapati juga bahawa perisian yang sedia ada tidak dapat menapis halaman yang tidak sesuai ini dengan efektif dan tepat. Dalam satu kajian yang dijalankan oleh Hunter (2000) terhadap 4 perisian penapisan yang popular iaitu CYBERSitter, Cyber Patrol, Net Nanny dan Surf Watch, ia mendapati prestasi keseluruhan tidak efektif sepenuhnya kerana ia gagal menapis 25 % bahan yang tidak sepatutnya tetapi pada masa yang sama ia menahan 21 % bahan yang boleh diterima.

Berdasarkan kajian terhadap perisian yang sedia ada juga, didapati kesemuanya lebih menumpukan kepada bahan berbahasa Inggeris sahaja sedangkan bahan-bahan yang tidak sesuai untuk kanak-kanak dan remaja juga terdapat dalam bahasa lain seperti Bahasa Melayu. Oleh itu, perisian bagi menapis bahan-bahan tersebut perlu dihasilkan.

Kajian ini akan menggunakan pendekatan berasaskan agen dan mekanisma gabungan beberapa kata kunci Bahasa Melayu dalam membangunkan perisian tapisan internet bagi membolehkan ia berfungsi dengan lebih berkesan. Perisian berasaskan agen ini berkeupayaan untuk memeriksa kandungan halaman web yang cuba dilayari atau yang diiklankan dan sekiranya didapati halaman itu mengandungi unsur yang tidak sesuai, ia akan menghalang halaman tersebut daripada terlepas kepada pengguna.

MEKANISMA PENAPISAN KANDUNGAN INTERNET

Mengikut laporan oleh Center for Media Education (1999), terdapat 4 mekanisma yang paling banyak digunakan oleh perisian penapisan iaitu senarai hitam (*blacklist*), senarai putih (*whitelist*), kata kunci (*keyword*) dan sistem pengkelasan (*rating system*).

- Senarai hitam tidak membenarkan capaian bagi halaman yang terdapat dalam senarainya.
- Senarai putih hanya membenarkan capaian bagi halaman yang ada di dalam senarainya sahaja.
- Mekanisma kata kunci adalah dengan cara memeriksa teks dalam tajuk atau kandungan halaman web dan halaman yang mengandungi kata kunci yang ditetapkan tadi akan disekat.
- Kaedah sistem pengkelasan membenarkan pemilik halaman web untuk menetapkan sendiri tahap kandungan halaman tersebut berdasarkan beberapa kategori seperti bahasa, seks dan keganasan dengan setiap kategori dilabelkan mengikut tahap (0,1,2,3,4). Di sini perisian penapisan akan menapis halaman mengikut tahap yang dibenarkan.

Berdasarkan kajian yang dikendalikan dalam Projek Penilaian Penapisan Internet (TIFAP) dari April hingga September 1997, ia mendapati setiap mekanisma yang digunakan untuk menapis kandungan Internet mempunyai beberapa kekurangan iaitu :

- Mekanisma senarai hitam dan putih hanya bergantung kepada penilaian kandungan internet yang dilakukan secara manual oleh pengeluar perisian tersebut. Bagi kaedah ini, sekumpulan pekerja akan membuat pencarian ke atas laman-laman web yang dilarang atau dibenarkan . Kaedah ini boleh mengakibatkan penilaian yang tidak adil dan perlukan sumber tenaga yang akan mengemaskini senarai tersebut dari semasa ke semasa.
- Bagi mekanisma pengkelasan, hanya 4% daripada jumlah keseluruhan halaman web yang menggunakan mekanisma tersebut. Isu kebolehpercayaan akan wujud apabila pembangun halaman sendiri dibenarkan untuk melabel halaman mereka.
- Mekanisma kata kunci didapati tidak dapat menapis dengan baik kerana ia akan menghalang halaman yang mengandungi walaupun satu daripada kata kunci tersebut. Ia merupakan kaedah yang digunakan di permulaan carian dan kebanyakan sistem yang menggunakan kaedah ini menghasilkan reputasi yang buruk. Halaman yang disekat menggunakan mekanisma katakunci tidak boleh dilaksanakan tanpa mengecualikan sebahagian daripada perkataan *non-pornographic* ataupun yang seakan-akan perkataan *pornographic*. Contohnya perkataan seperti *'breast'* akan menghalang capaian ke laman *'breast cancer'* dan *'sex'* akan menghalang capaian ke laman *'Ann Sexton'*. Maka ini akan menyekat penyebaran maklumat kepada pengguna yang ingin mencari bahan yang berkaitan dengan perkataan yang ditapis.

SKOP KAJIAN

Skop kajian yang dibangunkan ini adalah menumpu kepada penggunaan agen iaitu perisian yang melakukan tugas bagi pihak suatu badan ataupun individu. Agen yang dihasilkan bertindak menggunakan mekanisma katakunci. Agen akan memeriksa teks dalam tajuk serta kandungan halaman web dan jika ia mengesan terdapat katakunci terlarang, maka halaman tersebut akan disekat. Katakunci yang digunakan dalam penentuan menapis halaman dipecahkan kepada dua iaitu penggunaan katakunci 'halang' dan penggunaan katakunci 'lepas'.

Penggunaan Katakunci 'Halang'

Dalam skop penggunaan katakunci 'halang' atau katakunci terlarang, agen yang dihasilkan mampu memantau dan menyekat mana-mana halaman web yang mengandungi katakunci katakunci yang termasuk dalam kategori tersebut. Agen tersebut akan membaca keseluruhan teks halaman web dan kemudian membandingkannya dengan perkataan yang terkandung dalam senarai katakunci 'halang'. Senarai katakunci 'halang' ini membantu agen untuk menilai samada halaman web itu perlu disekat ataupun tidak sebelum katakunci senarai 'lepas' menjalankan fungsinya.

Penggunaan Katakunci 'Lepas'

Dalam penggunaan katakunci 'lepas' atau katakunci yang dibenarkan, agen yang dihasilkan dapat bertindak dengan lebih efektif dalam pemantauan ke atas kandungan halaman-halaman web. Senarai katakunci 'lepas' mengandungi katakunci katakunci yang dibenarkan dan akan bertindak bersama-sama dengan senarai katakunci 'halang' dalam menentukan samada halaman web itu perlu disekat atau tidak dengan melakukan perbandingan selepas perbandingan senarai katakunci 'halang' dilakukan. Ini bermakna sekiranya halaman web itu mengandungi katakunci 'halang', agen akan merujuk kepada senarai katakunci 'lepas' untuk membuat keputusan samada untuk membenarkan capaian terhadap halaman tersebut atau menyekatnya.

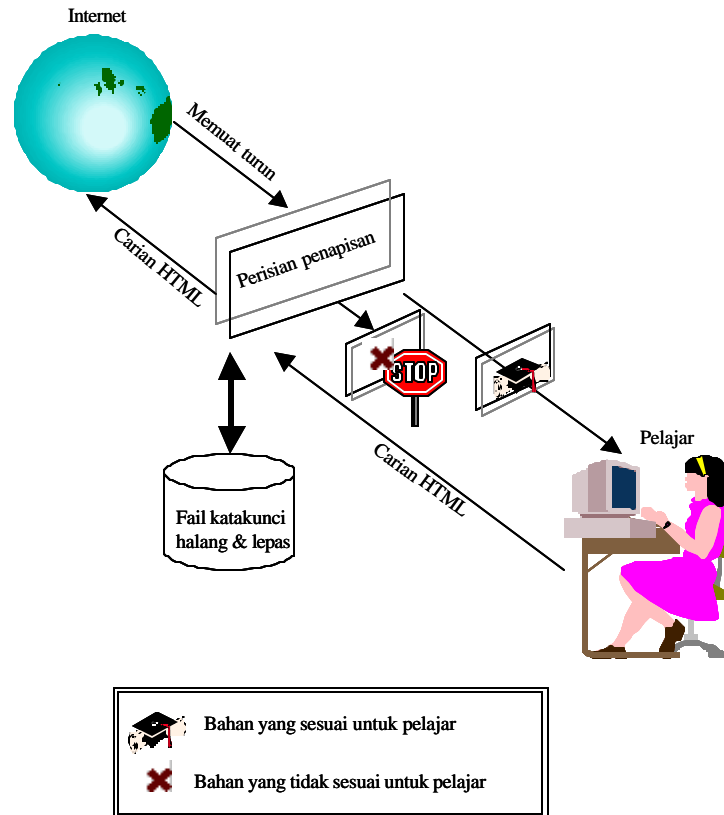
METODOLOGI KAJIAN

Dalam membantu membangunkan sesebuah aplikasi, terdapat beberapa kaedah yang dapat digunakan. Bagi tujuan membangunkan prototaip agen penapisan kandungan Internet ini, pendekatan yang telah dipilih ialah pendekatan atas bawah (*top-down approach*) yang juga dikenali sebagai Kitaran Hayat Pembangunan Sistem (*System Development Life Cycle – SDLC*). SDLC digunakan kerana ia menggunakan pendekatan penyelesaian masalah dan menetapkan satu piawai serta memerlukan laporan dari semasa ke semasa.

Dalam peringkat pembangunan sistem, perisian ini menggunakan teknik prototaip yang boleh menapis kandungan internet dengan menggunakan mekanisma katakunci. Prototaip ini mengaplikasikan ciri-ciri agen untuk membantu pengguna dalam menapis halaman yang mengandungi katakunci tersebut. Secara amnya, perisian berasaskan agen ini akan mempunyai keupayaan seperti berikut :

- i. Beroperasi secara sendiri (autonomous) tanpa penglibatan pengguna.
- ii. Mampu berkomunikasi dengan pelayar internet dan halaman web.
- i. Berperanan untuk menilai halaman web yang hendak dicapai dan membuat keputusan untuk menyekat atau melepaskannya kepada pengguna.

Dalam sistem ini, agen penapisan akan berinteraksi dengan Internet dengan melakukan carian URL seperti yang diminta oleh pengguna. Kemudian agen akan melihat keseluruhan dokumen dalam halaman web yang dicapai serta membandingkan setiap perkataan dalam dokumen tersebut dengan katakunci yang terdapat dalam fail senarai halang dan lepas. Bagi halaman web yang mengandungi bahan-bahan yang sesuai untuk pelajar, agen ini akan melepaskannya untuk tatapan pengguna sebaliknya jika tidak sesuai ia akan menyekat daripada dicapai oleh pengguna tersebut. Struktur perisian penapisan Internet yang berasaskan agen ini ditunjukkan seperti dalam Rajah 1.



Rajah 1: Struktur Perisian Penapisan Internet

KEPENTINGAN KAJIAN

Terdapat beberapa faedah dan kelebihan yang dapat diperolehi daripada pembangunan projek ini. Antaranya ialah:

- i. Sistem ini penting dalam membantu pihak kerajaan dan masyarakat terutama sekali bagi ibu bapa supaya anak-anak mereka tidak terdedah kepada unsur-unsur negatif yang terdapat dalam Internet.
- ii. Membantu organisasi dalam membuka laluan yang lebih luas kepada penghasilan dan pembangunan sistem penapisan seperti agen ini.
- iii. Menyumbang kepada limitasi dalam penyebaran maklumat dari Internet yang selama ini tidak terhad kepada apa jua bahan.
- iv. Menyumbang kepada sekatan bahan-bahan terlarang yang diwujudkan dalam Bahasa Melayu.

CADANGAN

Sistem penapisan yang dibangunkan ini hanya merupakan satu prototaip yang digunakan untuk tujuan menguji keberkesanan penggunaan dua katakunci. Bagi meningkatkan keberkesanan dan keupayaan sistem ini pada masa hadapan, beberapa cadangan perlu diambil kira iaitu memasukkan sedikit unsur kepintaran agar peratusan ketepatannya lebih tinggi dan

juga boleh ditambah dengan kemampuan untuk menapis halaman web yang mengandungi gambar-gambar lucu.

RUJUKAN

Beale, R., Wood, A., "Agent-based Interaction", *Proceedings of HCI'94*, Cambridge University Press, pp. 239-245.

Center for Media Education, "Children and Media", 2000, <http://www.cme.org>.

Center for Media Education, "Youth Access to Alcohol and Tobacco Web Marketing: The Filtering and Rating Debate", 1999.

Clinton, W. *Remarks by the President at event on the e-chip for the internet*, The White House Office of the Press Secretary, 16 July 1997.

Hunter, C. D., "Negotiating the Global Internet Rating and Filtering System: Opposing Views of the Bertelsmann Foundation's Self-regulation of Internet Content Proposal", *Computers, Freedom, and Privacy 2000 Conference (Toronto)*, April 2000.

Hunter, C. D., "Internet Filter Effectiveness : Testing Over and Underinclusive Blocking Decisions of Four Popular Filters", *Social Science Computer Review*, Vol. 18, No. 2, Summer 2000.

Lieberman, H., "Autonomous Interface Agents", *ACM Conference on Human-Computer Interface (CHI-97)*, Atlanta, March 1997.

Michael, W., Nicholas, R. J., "Intelligent Agents: Theory and Practice", *Knowledge Engineering Review*, January 1995.

Turow, J., "The Internet and the Family: The View from parents the view from the press", Annenberg Public Policy Center of the University of Pennsylvania, Report No. 27.