
A looking-out portal (LOP) approach to enhance qualitative aspects of bandwidth utilisation in academic networks

Waseem Ahmed* and Zahid Ansari

P.A. College of Engineering,
Mangalore 574 153, India

E-mail: waseem@computer.org E-mail: zahid.ansari@acm.org

*Corresponding author

Johannes Herrmann

Curtin University of Technology,
G.P.O. Box U1987 Perth, 6845, Western Australia
E-mail: jherrmann@ieee.org

Matin Abdullah

BRAC University,
Dhaka 1212, Bangladesh
E-mail: mabdullah@bracuniversity.net

Abstract: Campuses of educational institutions periodically need to increase network bandwidth to keep up with increased demand and this decision is based on the quantitative aspects of the network bandwidth utilisation. The qualitative utilisation of the bandwidth is seldom looked into. Improving the qualitative utilisation of the bandwidth may not even necessitate a network upgrade. Although blacklist-based access control techniques help to a certain degree, the findings of this research indicate otherwise. A multi-tier, whitelist-based, looking-out portal (LOP) approach is presented that promises to improve the qualitative utilisation of the network while positively impacting pertinent resource identification and location of sources on the internet. The authors draw on their years of experience serving as students and staff in various campuses of universities and colleges in various countries while making recommendations.

Keywords: network bandwidth utilisation; network access control; academic network; web portal; search engines; looking-out portals; LOP; qualitative network utilisation; virtual organisations.

Reference to this paper should be made as follows: Ahmed, W., Ansari, Z., Herrmann, J. and Abdullah, M. (xxxx) 'A looking-out portal (LOP) approach to enhance qualitative aspects of bandwidth utilisation in academic networks', *Int. J. Networking and Virtual Organisations*, Vol. X, No. Y, pp.000-000.

Biographical notes: Waseem Ahmed is a Professor in CSE at P.A. College of Engineering, Mangalore. He obtained his BE from RVCE, Bangalore, MS from the University of Houston and PhD from the Curtin University of Technology, Western Australia. His current research interests include

multicore/multiprocessor development for HPC and embedded systems, and computers in education. He has been exposed to academic/work environments in the USA, UAE, Malaysia, Australia and India where he has worked for more than a decade. He is a member of the IEEE.

Zahid A. Ansari is a PhD candidate in the Department of CSE, Jawaharlal Nehru Technical University. His areas of research include data mining, databases and MDD. He received his ME from Birla Institute of Technology, Pilani. He has worked at Tata Consultancy Services (TCS) where he was involved in the development of cutting edge tools in the field of model driven software development. He is currently with the P.A. College of Engineering, Mangalore as a Faculty. He is also a student member of ACM.

Johannes U. Herrmann is a PhD student at Curtin University of Technology, Western Australia where he is a member of the Institute for Multi-Sensor Processing and Content Analysis (IMPCA). He received his BSc (Hons., Mathematics and Computer Science) and MSc (Computer Science) from the University of Western Australia. His research interests include network reliability, artificial intelligence in online games, program synthesis and computers in education. He is a member of the IEEE.

Matin Saad Abdullah received his BSc in Electrical Engineering from the University of Houston and MS in Computer Science from the same institution. He has more than ten years of experience in the software industry and is currently teaching at BRAC University, Bangladesh, as a Senior Lecturer.

1 Introduction

The pervasiveness of computers and the internet in educational institutions worldwide is undeniable. Institutions strive to offer the best of these facilities to the academic community where possible. The proliferation of WiFi enabled mobile devices on campus has further contributed to the dependency on the internet. Services offered to the student community range from subject related material being made available on learning management systems (LMSs) to personal e-mail addresses and web spaces to the availability of fees, grade and other registration details being made online.

While these services seek to increase the academic value and convenience of the community within the campus, the internet, like other technologies, has a side that brings harm to these same educational values. This problem of a technology being used for both benefit and harm is referred to as the *dual-use dilemma* as further elaborated in Rashid et al. (2009). The article mentions that “The issue is not, in most cases, a call for a moratorium on the development of technology but rather an attempt to find ways of maintaining the benefits while avoiding or minimizing the harm.”

Campuses in many countries today are, in fact, addressing both these aspects. By periodically increasing the bandwidth of the campus network and by increasing the number of computers on campus with access to the internet the educational institutions are trying to maximise or maintain the benefit the internet provides. And by enabling various access control techniques on the internet traffic and implementing download limits the institutions try to minimise the harmful effects. To cite an example, in a developing country like Bangladesh this problem is compounded by the cost of internet

access for households. Since both on one hand, the cost of the internet service is relatively high and on the other hand, the purchasing power of the households are relatively limited the university internet connection is abused the most. In one of the Universities in Bangladesh, where one of the authors taught, on top of blacklist-based access control everything has been tried from using a different proxy for the students to allocation of time period for Facebook access. However, the *deperimeterisation* of campuses, as referred to in Pieters and van Cleeff (2009), has led to the disappearance of boundaries between the campus and the external world. This deperimeterisation raises many issues and questions not only on the security front but also on the fading accountability of the users within campus.

While there have been various studies on the beneficial and harmful effects of the internet on students, this paper seeks to address the effective utilisation of the network bandwidth in an academic setting. The authors draw on their years of experience serving as students and staff in various campuses of educational institutions in India, Australia, Malaysia, United Arab Emirates and the USA while making recommendations.

2 Related research

There has been much research on different aspects of the internet, its usage and influence on society based on geographic areas as given in Mikami (2002), Lee and Zhu (2002), and Veenhof (2006) and based on user behaviours as given in Nie and Hillygus (2002), DeBell and Chapman (2006), Mokhtari et al. (2009) and Shklovski et al. (2006). While the positive aspects of the internet and its strong influence on education are undeniable, the negative influences are also as many contributing to the dual-use dilemma of the internet.

Three cases taken to illustrate the dual-use dilemma of Google maps, chat rooms, peer-to-peer (P2P) networks, and social networking sites by Rashid et al. (2009) highlight both aspects of the internet. The negative impact of the internet on reading habits has been studied in Mokhtari et al. (2009). Anecdotal and other experience suggests that written communication skills of habituated users of sites like Twitter have resulted in language deterioration particularly among student populations who tend to communicate using language similar to that used in SMSs. The addictive nature of massive multiplayer online role-playing games (MMORPG) has an influence on student's learning and class attendance. *Character assassination*, as referred to on Wikipedia (2009), on social networking sites like Twitter, Facebook and Orkut also have adverse psychological impacts on individuals as given in TOI (2009). Rashid et al. (2009) report from (www.australian.news.com.au/story/0.24897.2532576212332.00.html) that 68% of students using Facebook had a significantly lower grade-point average than those who did not use the site. Karahasonovic et al. (2009) cite low-quality content (23%), harassment/bullying (9%) and time consuming (6%) as reasons for users stopping to use social networking sites.

From the network utilisation perspective, the campus bandwidth may be consumed by various non-academic related activities. Take e.g., the case of MMORPG, *mashups* and P2P networks. Although one student connecting to a MMORPG does not consume a large amount of network bandwidth as identified by Chen et al. (2006), and Griwodz and Halvorsen (2006), many students doing so at once can lead to a large amount of

bandwidth used over a period of time. Mashups, as referred to in Kendall (2007), are websites that combine data or services from two or more external sources, and cause much unwanted and unsolicited web traffic which may include spyware and adware or links to many of these. Also, content offered on mashups such as offers related to travel, hotels, education, rents, dating, car deals, etc. may be distractive and not directly relevant to the education of the student user. Hundreds or thousands of searches every second on Gnutella, a P2P network, relate to sexual content as given in Rashid et al. (2009).

The expectations of students to enjoy recreational computing on campus will no doubt continue to rise (Rickman et al., 2003). According to statistics cited in Rickman et al. (2003), 70% of college students play computer games at least once a day. Educational institutions in many countries across the world need to implement some form of mechanism like internet access control or prefetching Kim et al. (2000) to improve the network utilisation. The basic steps range from denying access to sites that serve adult content and sites related to movies, drugs, violence, politics and religion to sites that are known for their spyware, malware and virus content.

While these access control steps do improve the network utilisation to an extent, our findings suggests that these ways are far from being effective on campus networks. Additionally, the effect of the greedy consumption of the network bandwidth on campuses may have an impact well beyond academic networks affecting the whole internet as well (Shum et al., 2001).

3 Campus networks

3.1 Site classification

The management and the network administrators of many educational institutions are faced with the task of implementing some form of access control to the internet based on a blacklist. An important first step in achieving this is the binary classification of sites into allowable and not-allowable ones. Considering the existing size of the internet and its continuous growth, this is indeed a daunting task. Commercially available lists and those provided with access control software help ease this job.

Table 1 Site ranking based on academic relevance

<i>Cat.</i>	<i>Academic relevance</i>	<i>Types</i>	<i>Examples</i>	<i>Comments</i>
1	High	University sites, research related sites, technical forums	<i>Slashdot, Wikipedia, NASA, *.edu, sourceforge</i>	Educative
2	Medium-low	News and personal e-mails	<i>Gmail, Rediff, YouTube, CNN, BBC</i>	Recreational/informative
3	Low-very low	Social networking sites, downloading sites, P2P networks	<i>Facebook, Orkut, RapidShare, BitTorrent, EDonkey</i>	
4	Very low-Nil	Adult, racial, drugs, games	<i>SecondLife, WorldOfWarcraft</i>	Disruptive

The classification of a website as useful or relevant to a student from an academic perspective is both subjective and controversial. This article takes a conservative view and uses the classification of websites as given in Table 1 for illustration. The sites are ranked from 1 to 4 with sites in category 1 referring to websites that serve content of high academic value to the student and sites in category 4 referring to websites that serve content of very low or zero academic importance to the student. There may be exceptions to this classification. For example, a student majoring in game programming may be required to visit sites in category 4 as part of his course or a graduate student who needs to discuss and share his research findings with his other geographically separated peers on social networking sites in category 3. The classification does not consider such exceptions and is based on the general student population.

Also, Table 1 does not classify the highly used search engine portals like Google, Yahoo and Scroogle. The reason is that these portals are not direct sources of information but rather a means by which the student user is directed to other sources of information. The links on these portals, thus, may point to sites in any category based on the search terms entered by the user. The table also does not classify mashups as the content on them is dynamic and may fall into multiple categories.

3.2 Access control techniques

Two popular open source-based tools used with the Squid (2009) proxy server are DansGuardian (2009) and SquidGuard (2009). While these tools help in checking improper bandwidth utilisation to a large extent, experience and findings in this paper suggest that they are far from being completely effective on college campuses.

DansGuardian (2009) is a true web page content filter that uses combinations of methods like phrase matching, PICS filtering, MIME types and URL filtering to decide to allow access to web content.

SquidGuard (2009), on the other hand, is a URL redirector that is based on a blacklist. It uses a variation of the Aho-Corasick algorithm (Aho and Corasick, 1975), a string-pattern matching algorithm, for fast and efficient searches. As the search time is not dependent on the size of the blacklist the blacklist can be generously populated. Sites such as Shall Secure Services (2009) offer frequently updated manually categorised blacklists for free downloads to educational institutions and contain over 1.6 million entries.

Another approach used at some campuses is the partitioning of bandwidth between faculties, students on campus and students at the hostels by using separate proxies with separate access control mechanisms.

From the network management perspective, costs incurred in managing, detecting and dealing with problems associated with recreational computing continue to increase (Rickman et al., 2003). On some campuses, whitelist-based access has been imposed as a security measure. One such example is the case study at the TAMU campus given in Safford et al. (1993). Initially, a network intrusion was observed, monitored and blocked successfully at TAMU. A breach three days later after a complete cleanup and fixing of the network was followed by more successful break-ins. This could be contained only after changing the network filter which replaced deny-based filtering (let everything in unless it is specifically denied) to allow based filtering (block everything unless it is

specifically allowed). The article reports that no successful attacks were reported after this replacement (Safford et al., 1993).

3.3 Usage patterns

One of the greatest challenges for the information systems department is framing, enforcing and monitoring acceptable use policies of academic networks. These policies need to be framed comprehensively to encompass the various academic benefits available from the proper use of the network while limiting its recreational use. For example, although Northwest Missouri State University's campus provides PCs in every residence hall room for the benefit of the students, it had been observed that students often bring gaming equipment such as X-boxes to campus (Rickman et al., 2003).

To identify network access patterns on campus networks, Squid Proxy access log files from two separate campuses were studied. IP addresses and user names were randomly obfuscated in the access log files prior to analysis to protect the privacy of users. Both campuses implemented some form of access control with one using DansGuardian and the other SquidGuard. Squid served as the proxy server in both cases. The campuses differed in other ways as well – one campus allowed access only during regular college hours whereas the other was an always on connection. One imposed accountability which required users to login with their individual user names and passwords while the other allowed anonymous login.

An analysis of the Squid access log files using the *Squint* (2009) log file analyser revealed the following trends in a three-week access pattern in a mid-semester period for the two campuses. Users were ranked based on the number of pages downloaded, number of sites visited and the number of bytes downloaded. It was noted that counts of some of the top ranked users in the latter two categories were 800%–1,000% above the counts of the average user. Users ranked high on the number of sites accessed stood about 400%–500% more than that of the average user. Users frequenting social networking sites like Facebook and Orkut accounted for the high number of sites accessed. Top sites visited for many average ranked users were Gmail, Rediff and YouTube that belonged to category 2. Top sites visited for many low ranked users were research databases, university sites and Gmail, which fell into categories 1 and 2.

Using the classification given in Table 1, it was noted that more than 90% of the downloads belonged to categories 2 and 3. It was observed that the maximum bandwidth of academic networks was used for accessing a combination of sites in categories 3 and 4, with category 1 being minimally accessed. Although many of the sites belonging to category 4 were blocked, it was observed that some of the heavily downloaded links either contained references that clearly fell into category 4, indicating the incompleteness of the blacklist, and/or contained material that should have been classified as category 4. As sites and downloads had to be manually inspected to classify them as educational, recreational, prohibited or strictly prohibited an orthodox view was taken to ensure consistency.

A market-basket analysis to predict relationships such as “users accessing site A also access sites B, C and D” using basic data mining tools (Alevoor et al., 2009) seemed relevant but posed a different challenge. Squid access logs contain the complete URLs of every site or link accessed which includes sites and links that were intentionally accessed as well as those inadvertently linked. For example, links to logos or product pictures may appear multiple times on pages at different levels of a company site and are thus are

recorded multiple times in the access logs. Another similar case is the advertisements from external sites on mashups. The link or site count in the access logs, thus, may not serve as a proper input for a market-basket analysis type of prediction.

The findings above are for campuses where access control was implemented. In campuses where such controls were not implemented, the statistics may be very different.

3.4 Limitations of blacklist-based access control

DansGuardian, although a very popular access control mechanism, has the following limitations

- 1 The entirety of genuine sites may be blocked as a result of phrase matching.
- 2 Specific genuine URLs are denied access for the same reason as above. In addition, genuine phrases such as ‘breast feeding and sex after pregnancy’ entered as queries to search engines like Google and Yahoo may be denied access.
- 3 Denying access to downloads with certain MIME types like *mp4*, *mpeg*, *flv* and *dvi* and to sites like YouTube that serve such content cannot be generalised as videos, movies, lectures and podcasts may actually be educational in nature.
- 4 DansGuardian assumes that adult sites strictly adhere to the PICS rating (Resnick and Miller, 1996) which, in reality, is not the case. This prevents many illegal sites from getting blocked.

For SquidGuard to be effective, the blacklist needs to be regularly updated and kept complete and correct. Considering the rapid expansion of the internet, documenting an appropriately categorised complete and correct list is far from being practically achievable.

Another disadvantage with the blacklist approach is the blocking of entire domains belonging to certain categories; e.g., the downloads category of the Shalla blacklist contains links to YouTube and Rapidshare which blocks the entire domain. While a large portion of these sites do belong to the non-educative category, there may be exceptions that indeed contain links to educational videos. Locating and subsequently verifying these links may be difficult if the domain is blacklisted.

Additionally, some sites featured on the blacklists use domain aliasing and site mirroring to good effect to enable them to fall outside the list. In addition, dedicated users on campus creatively use search engines like Yahoo and Google to locate these sites outside of the blacklists. Additionally, users on campus are known to circumvent these measures in entirety by using mechanisms like external proxies and tunnelling. These latter techniques, however, can easily be detected and curtailed by vigilant network administrators.

4 Unassisted resource location and identification

4.1 Generic search engines

The World Wide Web can be seen as an information gathering and learning tool. As the web continues to grow, searching and locating relevant and complete information on a

particular topic from among more than a billion documents in an efficient manner becomes challenging. Lazonder et al. (2000) goes to the extent of claiming that the “chance of instantly getting a list of useful references is close to zero”. Although search engines like Google and Yahoo have made it less difficult, extracting the proper documents from the thousands of matches turned up by a search engine in response to a query still remains laborious. To find the best match, users have to read through all the retrieved documents and identify the information manually. This makes the task of searching, accessing, extracting, interpreting and processing difficult and time-consuming (Lausen et al., 2004). The search process largely depends on the domain knowledge of the users and their past experience (Fidel et al., 1999). In their article, Lazonder et al. (2000) explain that subject teachers, rather than librarians have the responsibility of increasing the student’s level of domain expertise; this expertise is essential for producing a greater number of correct solutions in the search process while consuming lesser time.

Adequate training and a system design that can appropriately complement a user’s information seeking and searching behaviour are needed to realise the true potential of the web. The training should involve the evaluation of the results from the search process and should help in turning the search process into a productive and satisfying experience according to the study done by Fidel et al. (1999). The study also suggests that the productivity of the search would have further improved if spelling and typographical errors were corrected retrieval of sites with partial or incomplete URLs allowed and the filtering of non-useful sites. Although some of them are present to a certain degree, many of these features are absent from the commonly used search engines.

4.2 Domain specific web portals

Web portals, as described by Lausen et al. (2004), are entry points over the internet for a community of users that helps to present and exchange relevant information among that group of users. Pienaar (2003) describes web portals as customised gateways to information on the web. Highlighting the importance of the web, Akeroyd (2001) mentions that “building web portals, creating and adding value to web resources are all central to future educational systems”. The creation of a portal is thus an intense exercise which involves addressing usability, navigability and other aspects of HCI. A portal’s consistent look-and-feel aids in the easier navigation on the large information space for the users. Navigability, as mentioned by Lausen et al. (2004), is a basic functionality as proposed in the design guidelines for building effective websites.

The scholar’s portal approach suggested by Pienaar (2003) aims at improving the retrieving capabilities of search engines by providing highly focussed searches that direct users to dependable sources of information. The article also identifies the importance of allowing academics to evaluate and add information sources to portals.

The institution portals provide a well guided hand-held assistance to the external users to enable them to easily locate what they wish to see. Additionally, search engines provided on the websites provide users a constrained domain of search (limited to the university site). The creation of these often impressive portals involves large focus groups and teams and is a large and ongoing exercise (Coombs, 2007). Moreover, designing portals for specific age groups is a dedicated iterative process as has been described in Large et al. (2002).

5 Assisted resource location and identification

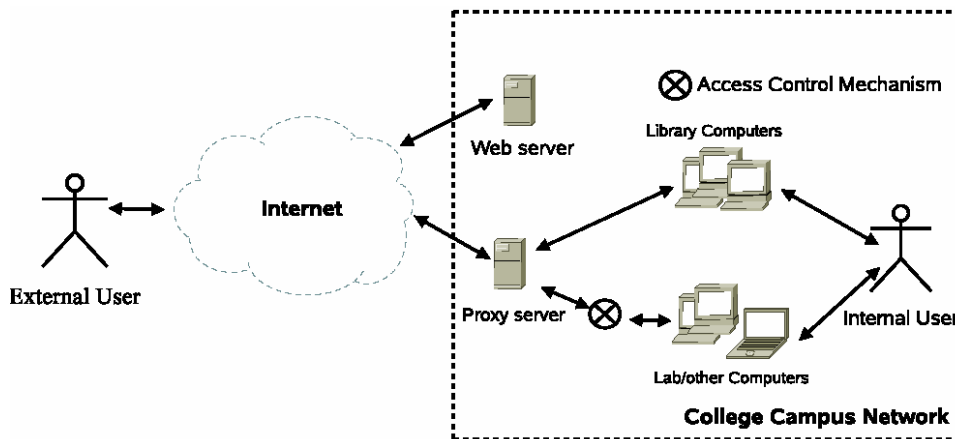
5.1 Ineffectiveness of existing methods

Consider a campus network depicted in Figure 1.

The external user on the left represents the category of users (generally from outside of the campus) 'looking in' to the university. This category of users may represent prospective students, parents of prospective students, staff from other institutions, alumni, curious visitors, companies looking into a university for campus recruitment, or other potential users. This category also represents current students accessing the web server through a proxy.

The other user category depicted on the right hand side of Figure 1 represents the users 'looking-out' into the vast cyberspace from within the college campus.

Figure 1 Looking-in and looking-out (see online version for colours)



For this user, a looking-out portal (LOP), corresponding to the hand-held guided approach of the institutional portals, referred to in this article as the looking-in portal (LIP), is generally absent. Users are mostly left to themselves to explore the vast cyberspace. The following situations may arise in this case

- 1 This exercise on the user's part may or may not turn up pertinent links depending on the creativity and experience of the user with the available search engines. A user without sufficient knowledge of the domain might look for one piece of information using a certain keyword or key phrase without realising that another equally important component of the domain not matched by the keyword or key phrase is also highly relevant (Schraefel, 2009). Results from a study conducted by Lazonder et al. (2000) reveal that experts with superior skills are more proficient at locating websites than those with less experience and that the search performance is enhanced by domain expertise. This is also attributed to the fact that a search process involves the asking of right questions, consulting proper sources of information and creatively combining the search outcomes to extract relevant information.

- 2 Also, when valuable domains or links are discovered, they are not systematically disseminated to others sharing their interest. In fact, if the collection of links in a sitting or the keywords or phrases used in the search is not 'preserved' by these users they may be lost as well with the users not being as successful in their subsequent search.
- 3 There is a marked absence of procedures for systematic link *retention* and *dissemination* although they are definitely important in the current educational system. Knowledge building by association, collections, history review, collaborative knowledge building and collaborative research referred to in Golovchinsky et al. (2009), Chi (2009) and Schraefel (2009) are all absent in the existing mechanism adopted by the existing 'looking out' type of search.

5.2 *Whitelist-based LOP*

LOPs are not new. Instructors and research groups in many institutions maintain subject specific portals that provide students with well guided 'paths' to the outside world which students, fellow researchers and colleagues find extremely valuable. These portals serve as both LIPs for the external users, as well as LOPs for the inside community. A LOP is a portal similar to a LIP that presents relevant and comprehensive appropriately categorised links on various subjects. It is a continuously grown portal that preserves educative and informative links for other users for subsequent access. This retention and dissemination of links on a portal helps a future user who may instead have to rely on a chance discovery in a search engine to locate that particular link. As Lazonder et al. (2000) point out, a search is basically executed to locate a particular website from which the relevant information is retrieved. If a link to this particular website had been made available to the user in some form, the web search process to locate that link would not have been necessary. In fact, appropriately structured query strings, efficient search engines and a systematic manual process of filtering from the sites retrieved by the search engine are essential to locate proper websites.

By narrowing down the search time for a range of users, a LOP can help reduce unnecessary wastage of network bandwidth. To further improve the bandwidth utilisation, a network administrator may choose to either cache a local copy on the proxy for frequent downloads or, if the copyright permits, to explicitly store a copy for offline and frequent use.

5.3 *Creation and growth of a LOP*

The creation of a LOP and its subsequent growth is an iterative process and needs to involve the LIS, different departments and the campus IT personnel among others. The LOP can be hierarchical but decentralised in structure for easier maintenance as described in Table 2. Brophy (2000) points out that responsibility of scoping collections at the research group or course level need not necessarily rest on library alone. Individual staff teaching courses and research groups within large organisations host their own LOPs. These fall under tier 3 and tier 4 respectively under the classification given in 2. Direct and focused links to both external content such as specific articles in trade and research journals and open source software; and internal content like lecture notes, staff research papers and product specification sheets. These links are extremely valuable in terms of

the reduction of search time, relevance, focus, and quality (authenticity). For example, staff at Curtin University of Technology's Bentley Campus in Perth has maintained very valuable links on many of their undergraduate course pages hosted on a LMSs. Newsgroups provided within the site also help students socialise and interact at the subject level (tier 3).

Departments within an institution need to work in conjunction with the LIS in the design of a LOP. Procedures need to be formulated so that links can be authorised and approved as educative by the concerned departments for inclusion in the LOP. The collection is gradually grown based on recommendations by the concerned staff of various departments. The growth of the LOP will mirror the growth of a library in an academic institution in many ways – a rapid growth curve initially until a critical mass is reached after which it becomes stable with less frequent changes. This is when a LOP has stabilised or matured. Until then, the readers do not have a proper comprehensive collection when looking-out. The collection is gradually grown based on recommendations by the concerned staff of various departments.

Table 2 Hierarchy of a LOP

<i>Tree level</i>	<i>Category</i>	<i>Examples</i>	<i>Time to mature</i>	<i>Collection growth/updates</i>	<i>People involved in maintenance</i>
Tier 0 (root)	NA	NA	NA	NA	LIS
Tier 1	Departments (major)	CSE, ECE	Very fast	Slow	LIS
Tier 2	Minor subject areas	Software engineering, DSP	Very fast	Slow	LIS, staff of departments
Tier 3	General topics within the subject	Requirements engineering, SE process models	Fast	Incremental	LIS, staff/students of Department
Tier 4	Specialised sub topics	Ethnography	Slow	Incremental	LIS, staff and research groups of the department

As Brophy (2000) suggests, to ensure a library collection is adequate there is still a need to prescribe boundaries even in the electronic context where resources need to be selected and evaluated at an organisational level. For links that are not currently available but are potentially useful are put on the LOP based on approval from the relevant departmental authorities similar to the formal procedures that exists for an inter-library loan (ILL). A proper procedure for periodic checking for updated, changed and broken links needs to be formulated. For browsing content outside the LOP, separate computers need to be provided in the library to allow unrestricted access to the cyberspace.

6 Limitations, implications and future work

Implementing the LOP on campus will require overcoming two challenges.

Firstly, although the presented LOP-based approach is a cost saving model from the academic institution's perspective, it would be incorrectly perceived as a 'restriction-of-freedom' model from the student's perspective. Working against this mindset would indeed be a big challenge. One way to gradually transition to this model is to introduce the LOP alongside the existing access model; either allowing the LOP to have the larger share of bandwidth while limiting the access outside the LOP to a lesser bandwidth or implementing the LOP-based model during regular office hours while retaining the existing model outside these hours. Another way is to section out the campus network based on usage with separate firewalls, proxies and LOPs for each of the section.

Secondly, the HR problem is paramount in this case. The effort in man-hour required to maintain an updated and relevant LOP on a subject may become huge for large institutions. Automatic LOP sharing between academic institutions would help reduce the effort required in LOP maintenance. Research on how this can be automated across multiple institutions would be an interesting field of study.

7 Reflections and conclusions

While networks continue to grow in terms of users, increased bandwidth will be in demand continuously. In networks spanning multiple academic and research institutions where expensive equipment are shared between researchers and academicians over the network as in Salzmann et al. (1999) and the GARUDA Network Initiative (2010) in India, and where large educational content is delivered over the network as in Rickman et al. (2003) and Shum et al. (2001), security and network bandwidth become very important. Institutions providing internet access are under pressure to regularly increase their bandwidth. Odlyzko (2003) claims a growth of close to 100% per year in the internet traffic. This implies that the academic institutions need to, at a minimum, increase their network bandwidth proportionally to sustain user experience. While a quantitative usage of the campus network bandwidth can be used to arrive at a quick decision, a qualitative look at the network traffic gives a more correct indication of what is actually needed. An analysis of network traffic patterns in campuses of two different universities located in different countries was presented.

The LOP-based approach presented in this paper proposes to improve the qualitative utilisation of the network bandwidth while enhancing a student's overall learning experience. The presented approach is not new in the true sense of the word and attempts to bring the approaches adopted by individuals and groups at large institutions into the forefront. Improving the quantitative usage of the bandwidth is also an explicit goal of the presented methodology.

How effective is the proposed methodology? To evaluate its efficacy; the methodology needs to be adopted and implemented by an institution for a fixed test period. But prior to risk implementing it, the management of the institution would want proven results on its efficaciousness. This presents itself as the classic chicken and egg problem.

The benefits of this methodology can be easily identified with and foreseen as this has the potential for reducing search time amongst students who are less knowledgeable and experienced in the domain. This also will help focus the network utilisation of academic networks to the core purpose of education.

References

- Aho, A.V. and Corasick, M.J. (1975) 'Efficient string matching: an aid to bibliographic search', *Communications of the ACM*, Vol. 18.
- Akeroyd, J. (2001) 'The future of academic libraries', *Aslib Proceedings*, Vol. 53, No. 3.
- Alevoor, A.K., Bolar, A., Ashwini, B. and Pais, C.M. (2009) 'Frequent pattern mining to discover user access patterns', Project report, P.A. College of Engineering, Mangalore, India.
- Brophy, P. (2000) 'Towards a generic model of information and library services in the information age', *Journal of Documentation*, pp.161–184.
- Chen, K-T., Huang, P. and Lei, C-L. (2006) 'Game traffic analysis: an MMORPG perspective', *Computer Networks*.
- Chi, E.H. (2009) 'Information seeking can be social', *IEEE Computer*.
- Coombs, K.A. (2007) 'Building a library web site on the pillars of web 2.0', Information Today, Inc., available at <http://infotoday.com/cilmag/jan07/Coombs.shtml>.
- DansGuardian (2009) Available at <http://dansguardian.org/>.
- DeBell, M. and Chapman, C. (2006) 'Computer and internet use by students in 2003-statistical analysis report', Technical report, Institute of Education Sciences, US Department of Education.
- Fidel, R., Davies, R.K., Douglass, M.H., Holder, J.K., Hopkins, C.J., Kushner, E.J., Miyagishima, B.K. and Toney, C.D. (1999) 'A visit to the information mall: web searching behavior of high school students', *Journal of the American Society for Information Science*.
- Golovchinsky, G., Qvarfordt, P. and Pickens, J. (2009) 'Collaborative information seeking', *IEEE Computer*.
- Griwodz, C. and Halvorsen, P. (2006) 'The fun of using tcp for an mmorpg', in *Proc. 2006 Int'l Workshop Network and Operating Systems Support for Digital Audio and Video Newport*.
- Initiative, G.T.N.G.C. (2010) Available at <http://www.garudaindia.in/>.
- Karahasonovic, A., Brandzaeg, P.B., Vanattenhoven, J., Lievens, B., Nielsen, K.T. and Pierson, J. (2009) 'Ensuring trust, privacy, and etiquette in web 2.0 applications', *IEEE Computer*, Vol. 42, No. 6.
- Kendall, K.E. (2007) 'Mashups: the art of creating new applications by combining two or more web sites', Decision Line, available at http://www.decisionsciences.org/DecisionLine/Vol38/38_2/dsi-dl38_2ecom.pdf.
- Kim, S-H., Kim, J-Y. and Hong, J.W. (2000) 'A statistical, batch, proxy-side web prefetching scheme for efficient internet bandwidth usage', in *Proc. Network and Interop Engrg. Conf.*
- Large, A., Beheshti, J. and Rahman, T. (2002) 'Design criteria for children's web portals: the users speak out', *Journal of the American Society for information Science and Technology*.
- Lausen, H., Stollberg, M., Hernandez, R.L., Ding, Y., Han, S-K. and Fensel, D. (2004) 'Semantic web portals – state of the art survey', Technical Report 2004-04-03, Digital Enterprise Research Institute (DERI).
- Lazonder, A.W., Biemans, H.J. and Wopereis, I.G. (2000) 'Differences between novice and experienced users in searching information on the World Wide Web', *Journal of the American Society for Information Science*.
- Lee, B.K.M. and Zhu, J.J.H. (2002) 'Internet use and sociability in mainland china and Hong Kong', *IT & Society*.

- Mikami, S. (2002) 'Internet use and sociability in Japan', *IT & Society*.
- Mokhtari, K., Reichard, C.A. and Gardner, A. (2009) 'The impact of internet and television use on the reading habits and practices of college students', *Journal of Adolescent and Adult Literacy*, Vol. 52, No. 7, pp.609–619.
- Nie, N.H. and Hillygus, D.S. (2002) 'The impact of internet use on sociability: time diary findings', *IT & Society*.
- Odlyzko, A.M. (2003) 'Internet traffic growth: sources and implications', in *Proceedings of ITCOM 2003*.
- Pienaar, H. (2003) 'Design and development of an academic portal', *Libri*, Vol. 53, pp.118–129.
- Pieters, W. and van Cleeff, A. (2009) 'The precautionary principle in a world of digital dependencies', *IEEE Computer*, Vol. 42, No.6.
- Rashid, A., Weckert, J. and Lucas, R. (2009) 'Software engineering ethics in a digital world', *IEEE Computer*, Vol. 42, No. 6.
- Resnick, P. and Miller, J. (1996) 'Pics: internet access controls without censorship', *Communications of the ACM*, Vol. 39, No. 10, pp.87–93.
- Rickman, J., Todd, K., Verbick, T. and Miller, M. (2003) 'The evolution of the electronic campus: From a communication system to an educational delivery system', in *SIGUCCS '03: Proceedings of the 31st Annual ACM SIGUCCS Conference on User Services*.
- Safford, D.R., Schales, D.L. and Hess, D.K. (1993) 'The tamu security package: an ongoing response to internet intruders in an academic environment', in *Usenix-ss'93: Proceedings of the 4th Conference on UNIX Security Symposium*, pp.1–56.
- Salzmann, C., Saucy, P., Gillet, D. and Mondada, F. (1999) 'Sharing of unique or expensive equipment for research and education', *Informatik*, Vol. 4, pp.32–43.
- Schraefel, M.C. (2009) 'Building knowledge: what's beyond keyword search', *IEEE Computer*.
- Shall Secure Services (2009) Available at <http://www.shallalist.de/>.
- Shklovski, I., Kiesler, S. and Kraut, R. (2006) 'The internet and social interaction: a meta-analysis and critique of studies', *Computers, Phones & the Internet: Domesticating Information Technology*, Oxford University Press.
- Shum, S.B., Marshall, S., Brier, J. and Evans, T. (2001) 'Lyceum: internet voice groupware for distance learning', in *Proceedings of Euro-CSCL 2001: 1st European Conference on Computer-Supported Collaborative Learning*.
- Squid (2009) Available at <http://www.squid-cache.org>.
- SquidGuard (2009) Available at <http://www.squidguard.org/>.
- Squint (2009) Available at <http://www.ledge.co.za/software/squint/>.
- TOI (2009) 'Caught in the net', available at <http://timesofindia.indiatimes.com/Mumbai/Caught-in-the-net/articleshow/4265381.cms>.
- Veenhof, B. (2006) 'The internet: is it changing the way Canadians spend their time?', Technical report, Science, Innovation and Electronic Information Division (SIEID), Ottawa.
- Wikipedia (2009) Available at <http://en.wikipedia.org/>.