



Viability and Utility of a Distributable and Searchable Cache System (DiSC) in a Developing Country

Omar Bakr, Aaron Chatterji, R.J. Honicky, Samir Mehta

November 22, 2004

Ghana



Abstract

This paper will investigate the impact of a Distributed Searchable Cache (DiSC) system on improving Internet access in the developing world. Our research has three objectives. First, we test the general viability of DiSC in a developing nation computing environment. Next, we examine the potential utility of DiSC in this setting, based on technical results, survey data, and interviews. Lastly, we analyze the potential drivers and barriers for adoption of DiSC. We find that DiSC is viable in a computing environment characterized by intermittent connectivity. We also find some evidence that DiSC will be useful in developing nations, because it improves document retrieval speed and reduces bandwidth demand. We argue that DiSC will be most beneficial in educational settings, and analyze the political, economic, social, and technological factors that will influence adoption.

Acknowledgements

This work was made possible by the generous support from the Management of Technology Program at UC Berkeley. The funding would not have been possible without the hard work of Professors Andrew Isaacs and Kristiana Raube who created this fellowship for conducting field studies in developing countries. Prior to our trip Ghana, both Richard Newton, Dean of the College of Engineering, and Eric Brewer, Professor of Computer Science, provided us with very useful insights regarding the design and development of DiSC and the conduct of our field study. Upon our arrival to Ghana, several people helped enrich our experience and make the process less painful. In particular, we are extremely grateful to Mark Davies the founder of BusyInternet in Accra. Mark provided us with very valuable information regarding the IT infrastructure in Ghana and the related political issues in that area. He also gave us access to the web proxy traces at BusyInternet. We could not have conducted the Internet usage surveys without the hard work of several students from the University of Ghana in Legon. In particular, we would like to thank Enunam, Sampson, Charles, and Barry. We would also like to thank Kwame from Calvin College, USA. We also owe a great debt of gratitude to Vic Dodoo, Barfi, Bernard and Anaba Alemna of the Balme Library at Legon, and Nina Marini and Ato of Ashesi University for giving us access to their computing facilities and allowing us to set up our servers. Several others provided us with good advice and support either prior or during the trip. These include Gregg Zachary, Eric Osiakwan, and Guido Sohne. We also received very good input from our fellow class UNIDO fellows both prior and after the trip. We are very thankful to all of them and anyone else whom we forgot to mention.

About the Authors



Omar Bakr is a second-year doctoral student at the department of Electrical Engineering and Computer Science at UC Berkeley. His research interests include distributed computing and wireless communications and networks. He was instrumental in the design of DiSC. He received both Bachelor and Master of Engineering degrees in computer science from the Massachusetts Institute of Technology.



Aaron Chatterji is a fourth-year doctoral student at the Haas School of Business at UC Berkeley. His research focuses on entrepreneurship, international business, management of technology and corporate social responsibility. He graduated from Cornell University with a degree in economics and international relations.



R.J. Honicky is a PhD student studying in the Department of Computer Science at UC Berkeley. Research interests include data distribution and management, and solving development problems using IT. R.J has worked as a software engineer at AirTouch Cellular (now owned by Verizon), Internet startups BabyCenter.com and Epsilon Inc., and storage vendor Network Appliance Inc.



Samir Mehta is a second-year MBA student at the Haas School of Business at UC Berkeley. He is interested in the process of technology transfer from industrialized societies to developing nations. He previously worked at eBay as a product manager and at Arthur D. Little as a strategy consultant. He graduated from UC Berkeley with a degree in Biochemistry.

1. Introduction

This paper discusses field research regarding the viability and utility of a Distributed Searchable Cache System (DiSC) in a developing nation computing environment with intermittent Internet connectivity. The software was motivated by frustration with the slow and unreliable connections in many developing nations. Our research objective is to investigate the viability and utility of DiSC in a developing nation setting. In addition, our field study aimed to gauge the demand for DiSC and investigate the potential obstacles to large-scale implementation.

Our approach is interdisciplinary and multi-method. We analyze the technical data from a pilot roll out of DiSC and report the results from approximately 200 user surveys and numerous detailed interviews with thought leaders. Our fieldwork was conducted in Ghana in May-June 2004, at several universities and Internet Cafes.

2. What is a Distributed Searchable Cache?

Caching is one of the most common and effective techniques in a computer scientist's toolbox. In the context of computer science, a cache is a (relatively) small but very fast place to store information that is frequently used. A cache allows a computer to avoid accessing a much larger (and hence more complete) but much slower data store. The more "locality of reference" (tendency towards reuse) in a data set, the better a cache performs.

Caching is frequently employed to improve performance of the World Wide Web. Virtually every web browser keeps a cache of recently used documents so that if and when that page is accessed again, it can be retrieved from the computer the web browser is running on, instead of requiring the computer to contact the web server. Similarly, organizations such as

corporations, schools and Internet Service Providers often utilize “proxy” caches, which intercept requests to a web server, and use their own stored copy of the requested document if possible.

Our software, which we call DiSC, is a proxy cache with some added functionality. The design of DiSC is based on the observation that a proxy cache can also be used to help “mask” (i.e. hide and mitigate the impact of) network outages. But there is another essential component to masking network outages: searching. Users more often than not navigate to information on the Web via a query to a search engine. Since search engines such as Google and Yahoo are not available during a network outage, this functionality must be replicated in a cache in order to effectively mitigate the effects of a network outage.

Finally, in order to reduce the cost of deploying a searchable cache, and to make our system robust to many different types of failure, our software takes advantage of the latent resources in a computer network. Whereas a typical proxy cache is located in a dedicated machine, our proxy cache is distributed throughout a local network of computers. This allows our software to utilize unused storage, memory and processing power on these machines in much the same way the peer-to-peer networks such as Gnutella, Napster, Kazaa etc. do so.

While none of these concepts are entirely new by themselves, the synthesis of caching, searching, and distributed storage creates a system which we believe is uniquely suited to addressing the problem of slow and intermittent Internet connectivity, particularly in educational environments.

3. Research Objectives

Our field research had three main objectives:

1. Test our assumptions about the network and computing environment in Ghana, namely that the network is intermittent, and that there is significant overlapping interest in content. We refer to this as testing DiSC’s *viability* throughout this paper, since DiSC was designed under these assumptions.

2. Determine the *utility* of DiSC in a developing nation computing environment based on user, system administrator, and bureaucratic perceptions.
3. Analyze the political, economic, social, and technological factors impacting Internet technology adoption in Ghana, specifically as it relates to DiSC.

We tackle Objective 1 by collecting and analyzing technical data, web traces and measures of network latency. Objectives 2 and 3 are addressed using evidence from surveys and semi-structured interviews. We apply a Political/Economic/Social/Technological (PEST) framework to examine the structural determinants of DiSC adoption.

4. Study Setting

4.1. Ghana

In testing the DiSC system, we hoped to target a developing nation with adequate basic infrastructure, well-educated citizenry, and sufficient computing resources. While DiSC could be successfully implemented in many nations, we were cognizant of the potential impact a pilot study could have in spurring later adoption and yielding large sample results.

Our research team chose Ghana as the project site because of its historical leadership in Africa, basic but rapidly developing telecommunications infrastructure, and its stable political climate. In addition, one member of our team had cultivated relationships with organizations during previous work in Ghana, which facilitated our research efforts.

Table 1: Ghana – Facts at a Glance

| | |
|--------------------------------------|------------------|
| <i>Population (July 2003)</i> | 20,467,747 |
| <i>Literacy</i> | 74.8% |
| <i>GDP (2002)</i> | US\$42.5 billion |
| <i>GDP Per Capita (2002 est.)</i> | \$2,100 |
| <i>Telephone (Main Lines) (2001)</i> | 240,000 |
| <i>Telephone (Cellular) (2001)</i> | 150,000 |
| <i>Internet Users (2002)</i> | 200,000 |

[<http://www.cia.gov/cia/publications/factbook/geos/gh.html>]

Ghana was the first West African nation to be connected to the Internet in 1994. [http://www.ghanaweb.com/GhanaHomePage/it/] The recent liberalization of the telecommunications sector has sparked the expansion of Internet access in Ghana. The basic Internet infrastructure clearly exists, as all of Ghana's government ministries, its parliament, 100 percent of international organizations, and 97 percent of all Non-Governmental Organizations (NGOs) are connected to the Internet. In contrast, few private sector firms and educational institutions are online due to high costs. [National Communications Authority of Ghana, 2000]

Accra, the capital of Ghana, had few Internet Cafes in 2000 and now has over 600. The average price to access the Internet is between \$0.75 and \$1.25 per hour, not a modest sum in a nation where per capita income is \$2000 (USD) [http://www.ghanaweb.com] it is clear that to increase access in Ghana, innovations must be made to lower access costs.

When Internet connections are set up, "connectivity becomes a problem because of erratic and slow connections". As one development worker put it, "just as interests in {Internet} services are increasing, we are faced with the problems caused by limited bandwidth in the country." [www.iicd.org]

Ghana's large, relatively well educated, English speaking population and its basic but underutilized infrastructure made it an ideal site for our research project. The gains from a low cost technological solution like DiSC, which can reduce the dependence on phone connections and increase document retrieval speeds, are significant for the development of human capital and private sector growth in Ghana.

4.2. *Data Sources*

Once we arrived in Ghana, we chose 3 primary sites for our pilot study: the campuses of the University of Ghana, Legon and Ashesi University in Accra, and BusyInternet Cafe. We hypothesized that our software would have maximum impact in a university setting due to a

hypothesized homogeneity of user searches, and a focus on research oriented activities. Using the computer equipment that we brought from the United States (which we later donated), we set up 6-networked computers where had been installed with DiSC. Our third location, BusyInternet, is the largest Internet Cafe in West Africa. We also conducted surveys at five other small Internet Cafes.

As luck would have it, the staff at the University of Ghana was on strike for all but three days of our stay, meaning that students were not given access to the library, and hence our equipment. While we were able to adapt to these circumstances to some extent, we did not have the same degree of control over the equipment at the other locations. As a result, while we were able to briefly install DiSC both at BusyInternet and Ashesi University, we were unable to gather the extensive records of people's web browsing behavior that we had planned to gather. Fortunately, the staff at BusyInternet graciously provided us with an excellent log from their own proxy server (a type of system closely related to DiSC). While the log does not contain all of the data we intended to gather, it does contain the majority of it.

Each location exposed DiSC to a different set of users. The university students and staff who did get an opportunity to utilize DiSC at Legon and Ashesi used the Internet for educational research and personal communication. BusyInternet had a much more diverse customer base for obvious reasons, and we expected more heterogeneity among users. We collected several technical results relating to performance and utility of DiSC. We also surveyed users of the DiSC system.

5. Study Design

Using surveys, interviews, and technical data, we aimed to triangulate our research questions, bridging qualitative and quantitative evidence. The technical data is useful to understanding network attributes and user search habits, but detailed surveys were necessary to understand user experience and preferences. In sum, the quantitative and qualitative

evidence were both used to assess the viability and utility of DiSC. Still, a broader question remained. Realizing that technology adoption is driven by economic, political, and sociological forces, we used detailed interviews with people we regard as “thought leaders”. These interviews yielded knowledge that will aid the further development of DiSC other technologies which might be appropriate for wider implementation in Ghana and other developing nations.

6. Data Collection and Methods

6.1. Technical Data

This section describes the data we gathered from computers at BusyInternet, Ashesi University and the University of Ghana.

6.1.1 Web Proxy Traces

A "trace" is an event log that records the behavior of an actor in a system. In our case, this means a log of web pages, images, movies, etc. that a web browser downloads when a person views a web page. A trace typically records the web page's address, its size, when the page was downloaded, which computer on the network downloaded it, etc. For privacy reasons, the raw trace data cannot be shared as part of this study. We will present results based on the aggregate of this data.

Because we were unable to gather data at Legon, our principle source of trace data is from BusyInternet. BusyInternet already uses a standard web proxy cache (without search functionality), which keeps a record of twenty-one days of activity for one hundred computers.

This data is particularly interesting for several reasons. First, it is a relatively large data set, which means that we can get statistically significant results from it, and in some cases even from small subsets of the data. Secondly, because BusyInternet is the most modern Internet cafe in Ghana, we can view it as an example of what Ghanaians (and possible Africans by

extension) would do with the Internet if the Internet were both responsive and relatively affordable.

On the flip side, this means that we are basing our conclusions on an atypical Internet cafe, and hence may not be drawing conclusions that are relevant to typical Ghanaians or Africans. Nonetheless, since we are studying how people *would* use the Internet if they had DiSC available to them, we believe that this is an interesting and relevant data set.

Proxy cache traces can be used to determine the effectiveness of our system under different configurations and scenarios, as well as to characterize users' interests and behaviors objectively.

5.1.2 *Network Statistics*

We have also continued to gather statistics on the operating characteristics of the Legon network. We have also configured computers at Ashesi University and BusyInternet to record these statistics, but have not yet received the data.

We gather three statistics from each site.

- 1) Latency
- 2) Bandwidth
- 3) Downtime

Latency is a measure of how long a message takes to travel from one place to another. Latency is an important metric in measuring the responsiveness of a system.

Bandwidth measures the rate at which data travels through a channel, in our case over a computer network. Bandwidth determines how long it will take to download large files.

Downtime is the amount of time the Internet is unavailable due to network outages. Outages occur because of computer crashes and in the case of the VSAT satellite connection at Legon, weather.

6.2. *Surveys*

User surveys were conducted at six Internet Cafes in the Accra area, varying in hardware, connection speeds, and access prices. The Cafes were located in different parts of the cities, and although we did not use a rigorous statistical sampling technique to decide on the locations, we believe they are reasonably representative of the hundreds of Internet Cafes in Accra. We discovered that many Ghanaians could not afford BusyInternet's prices, so they instead utilized cheaper Internet Cafes with dial-up connections. It was important to survey prospective users of DiSC, and those using slow connections are most likely to benefit from the technology. In addition to our team members, seven students from the University of Ghana, Legon helped to conduct the surveys. The survey consisted of basic demographic and socio-economic questions followed by detailed questions about Internet surfing patterns and preferences. Some of the most useful information we obtained is information on the sites most frequently from Internet users. In addition, we also collected detailed data on the amount of money spent on Internet and other professional services per week. The full survey instrument can be found in the appendix.

6.3. *Interviews*

In addition to the technical and survey data we collected, we also conducted ten interviews with "thought leaders" in the Ghanaian technology community. These interviews were conducted between May and June 2004, and consisted of 30-60 minutes of open-ended question and answer, depending on the subject's field of expertise. Interviews were scheduled using our contacts in the Ghana technology community and UNIDO, and were conducted in person, using tape recorders. We interviewed the Founder and the Business Manager of BusyInternet, ICT business and educational consultants, technology entrepreneurs, computer science professors, and government officials. By choosing leaders in the private and public sectors, we explicitly recognized the strong linkages between government and business in

developing nations. The knowledge gained through this process helped us to identify the key drivers and obstacles to widespread adoption of ICT, including concerns about technical education in Ghana, telecom infrastructure, and government regulation. The results from the interviews will be summarized in the “Principal Findings” section.

7. Principal Findings

We first present our technical results, which we believe illustrate the typically intermittent connectivity in Ghana, as well as some interesting information about how users at BusyInternet use the Web. Next, we explore our survey data to illustrate some common characteristics of Internet users. Then, using technical and survey results, we argue that while a large portion of information desired by users is dynamic, users frequently access static informational websites such as those relating to education and research.

Thus, we assert that DiSC would be most beneficial in educational settings, based on technical, survey, and interview data, and our assumption that data used in an educational setting tends to be more static. Finally, we use the knowledge gained in the expert interviews and outside research to analyze the potential for adoption, which we find to be promising.

7.1. Technical Analysis

This section discusses the data we gathered from computers at BusyInternet, Ashesi University and the University of Ghana. In order to determine the viability and utility of DiSC in a particular Internet setting, we have to consider the following key points. First, from the network perspective: How long does it take on average to deliver a Web page over the network? How often does the network become inaccessible? The latter can occur due to power outages or network failures (computer crashes or link failures). High latency and failure prone networks make caching more desirable, and the ability to search the cache more critical. Second, we need to examine the network from the usage perspective: How correlated are the usage

patterns of different users on the same network? What kind of content do these users typically access? Highly correlated access patterns may increase the effectiveness of caching and searching. Usage patterns alone, however, are not a very good indicator of the potential benefit of caching without any knowledge about how static the content is. DiSC works best when the content is static, valid for a significant period of time, and searchable. On the other hand, DiSC will not prove useful for managing communication-based dynamic information such as email and instant messaging.

7.1.1 Intermittent Connectivity

Internet Cafes and computer networks in Ghana connect to the Internet using a wide variety of technologies, ranging from standard telephone lines, to land based wireless connections, to satellite connections, to direct connection to the African fiber optic backbone (SAT3). Although some Internet Cafes are upgrading their connections to more reliable and higher speed types of connections, most Internet Cafes still use standard telephone based modems. Unfortunately, most of these technologies are subject to frequent outages due to power outages, weather, upstream network outages, and overload.

As a result, Internet access is intermittent and slow for most users. To make matters worse, Ghana lacks a local backbone, which prevents different ISPs from communicating directly with one another. As a result, there is little incentive to host content in Ghana, and hence most of the content is hosted abroad in Europe or North America. This means that when a satellite link is down due to weather, etc., the Internet becomes completely unusable. From the network standpoint, this suggests that most of Internet Cafes can greatly benefit from DiSC.

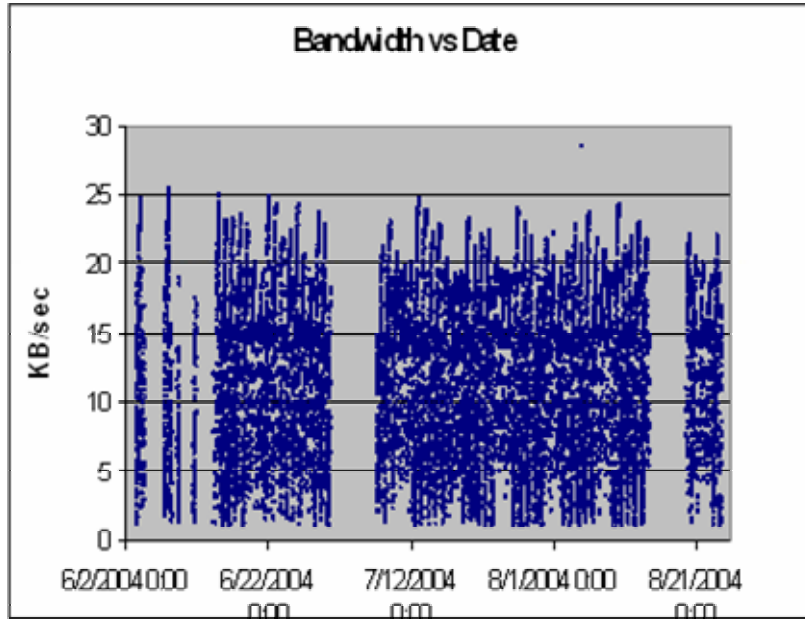


Figure 1: the bandwidth available to a single client in the University of Ghana network

In Figure 1, we plot the available bandwidth at the library of the University of Ghana at Legon for a period of approximately seven weeks. On the Y axis, we see the bandwidth available to a single client on the network, downloading a file from a server in the United States. On the X axis we see time. The gaps in the figure indicate downtime. The graph shows significant periods of downtime that can span days. While this data does not distinguish between internal outages that prevent the computer from accessing the gateway machine (which provides Internet access) on the one hand, and a satellite outage on the other, this distinction is not of primary importance for our purposes. The DiSC architecture is robust to both internal and external network failures, since a client will search and retrieve data from any other client it can access. If the machine is completely partitioned from every other computer, it can still search the data located on that particular machine. Using DiSC together with other techniques such as web page pre-fetching might significantly reduce the impact of these outages.

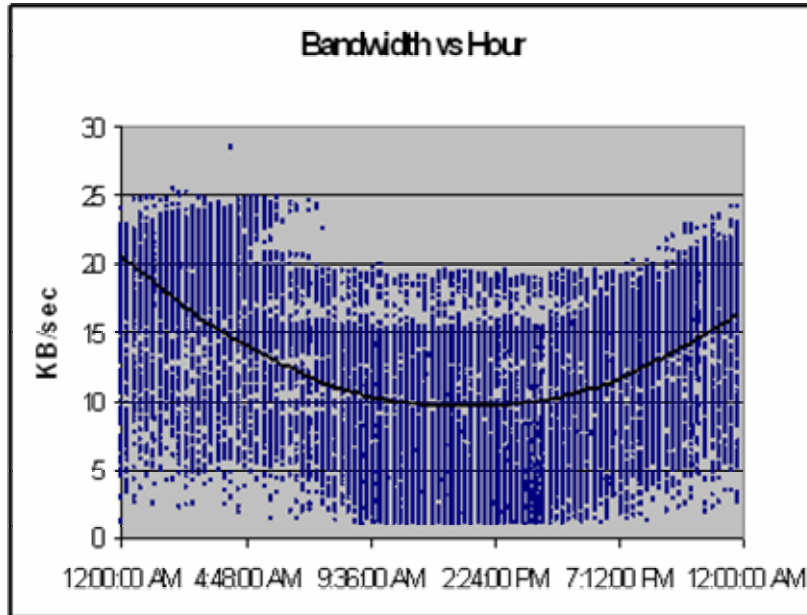


Figure 2: the bandwidth available to a single client in the University of Ghana network for a given time of day

Figure 2 plots the bandwidth available to a single client versus the time of the day, for the same period as above. From the figure, we can see that more bandwidth is available during the night hours. Obviously, this is partly because more users are competing for bandwidth during the day. If, however, this were the only explanation, we would expect to see high peak bandwidth even during the middle of the day, since there would be presumably little competition for bandwidth on weekends and holidays during the summer. Instead, we see that during the early morning hours, peak bandwidth is around 25KB/sec, whereas during the day peak bandwidth is around 20KB/sec. This suggests that daily weather patterns or sun exposure effect the bandwidth of the satellite connection.

Since the university pays a flat fee for the bandwidth from the satellite link, it makes sense to take advantage of these idle periods to do intelligent pre-fetching of web pages in order to mitigate the high bandwidth demands during peak hours and make the data available in case of a network outage.

7.1.3 Usage Behavior

Initially, we planned to deploy DiSC at a university library. This was based on the following assumptions: First, content fetched by library users is mostly scholarly work, which tends to be relatively static and searchable, and second, unique users can be grouped into different categories by common academic interests. Based on these assumptions, each user could potentially benefit from content fetched by others in the same category. For example, the larger the number of students who study biology, the higher the probability that they will find useful information by searching the cache. Such access patterns might not hold in other environments like Internet Cafes. In Internet Cafes, the scope of usage is much wider. In fact, the majority of users at the Internet Cafe we studied use the Internet to read e-mail and to chat with family and friends. For these purposes, DiSC does not add significant value.

Because the strike at the University of Ghana prevented us from testing these assumptions directly, we collected and analyzed data from two other sources. The first source is a survey of different users at various Internet Cafes. This provides a qualitative picture about the type of usage and content accessed. Survey results are discussed in the next session. The second source is a twenty-one-day web trace collected from a proxy server at BusyInternet. This provides exact statistics about websites and web services that were accessed during this period.

We had several observations from the data collected from the web trace. First, we looked at the correlation between queries to different search websites (e.g. Google, Yahoo, etc.). Our results show that there is very little overlap among search queries conducted on different days. These results suggest that if some of these queries were conducted over DiSC instead of Google or Yahoo, the likelihood of users finding meaningful results would be small because of the slim chances that the query had been issued before.

While this result might seem rather discouraging for DiSC, it is consistent with our expectations of users in Internet Cafes. Since we were unable to contrast the data from BusyInternet with data from the University of Ghana, we instead broke apart the data for some of the most frequently accessed domains, categorized them as personals, mail, news, info, music or travel (which were the most common uses of these domains). We then chose a representative domain for each of these categories and studied the hit rates of files for each of these domains.

| Domain | Hit Rate at “Knee” | Cache Size “Knee” | Total Cacheable Data |
|---------------------|--------------------|-------------------|----------------------|
| Americansingles.com | 92.5% | 4MB | 128MB |
| Cnn.com | 36.4% | 32MB | 64MB |
| Itafrica.com | 98.9% | 1MB | 2MB |
| Mtv.com | 31.2% | 64MB | 64MB |

Table 3: Hit rate for various “representative” domains

In Table 3, we see the document hit rates for four domains, plus the size of cache that was necessary to achieve that hit rate (beyond which there was not much benefit to adding more cache, hence labeled the “knee” of the curve in the table), and the total amount of cacheable data that was in the trace. While *americansingles.com* has a high hit rate, it was only necessary to cache a small amount of data to achieve that hit rate, suggesting that most of the hits were a result of inline data such as images, or default search results. Because the total cacheable data is so big for *americansingles.com*, we believe that this suggests that the corpus of relevant data is very large, and therefore difficult to cache and search. *CNN.com* shows a moderate hit rate that tapers off at around half the total corpus size, suggesting that a non-trivial amount of CNN’s data is “new” data, and that data from CNN is not relevant to BusyInternet users for a long time after it reaches the cache. *MTV.com* seems to have a large changing corpus that was not well explored in the two-week trace, and so it is difficult to speculate about how well DiSC would perform against site.

Most interesting is *itafrika.org*, which has a very high hit rate, a very small corpus, and is the 4th most popular sites in the traces. Furthermore, the data from *itafrika.com* is eminently cacheable. This is good news for DiSC, since in this case informational data is both easy to cache and repeatedly accessed. Since this is the type of data we expect to see in libraries, this is promising for DiSC. Thus, both search query analysis and an informal survey of the trace data suggest that DiSC might be useful in libraries that restrict their Internet usage to information retrieval.

7.2. Survey Analysis

After surveying 198 people from 6 Internet Cafes in Accra, we offer some preliminary conclusions that are useful to understanding the potential of DiSC and other related technologies.

Some descriptive statistics about our sample:

- 92% of our survey respondents had been to Internet Cafes before and 77% of respondents use the Internet at least once a week, so our sample represents relatively experienced and frequent Internet users.
- 71% of respondents were students and 84% between 18-33 years old.
- 73% of respondents were male.
- Our sample was disproportionately educated, with nearly 43% university educated or higher. While this percentage is much higher than Ghana as a whole, it is likely that Internet users tend to be more educated than the population as a whole, so this bias is not necessarily cause for concern.
- Only 20% of users had Internet access at home, meaning that a large majority of respondents use the Cafe as a primary location to do research or communicate with friends. However, 74% of respondents had access to a computer at home, perhaps suggesting huge potential for home Internet access in the future.

- Predictably, email, research, news, and chat dominated surfing preferences, while education and training was a surprisingly popular activity. This fact suggests that many Ghanaians are using the Internet to obtain skills not available to them locally. 89% have email accounts, and the majority has been using email for more than 3 years
- The most popular sites are Yahoo, Hotmail, Google, and two major Ghanaian sites, myjoyonline.com and ghanaweb.com.
- 20% of respondents operate their own business and 74% of business owners use the Internet Cafe for business.

These statistics offer us several useful pieces of information. First, the Internet users in Accra are educated and sophisticated users. We were very surprised by the number of respondents who have computers at home, and are suspicious that some individuals were concerned with how the survey team would view their socio-economic status if they truthfully reported that they did not own a computer. On the other hand, this may in fact be true, considering that refurbished PCs are relatively inexpensive.

Education and training activities were somewhat popular. This trend bodes well for the adoption of DiSC, since the software's natural advantage is in computing environments characterized by relatively homogenous users searching for static information. For example, research papers and training manuals could be easily and rapidly accessed using DiSC.

Finally, many entrepreneurs use Internet Cafes for business, a trend we had not previously anticipated. This empirical fact suggests future research should be conducted into what types of business activities are being done, and how DiSC may be useful.

The major finding from the technical and survey analyses of usage patterns and the users is that there is probably an opportunity to cache large amounts of static information. Our data indicates that users have common characteristics and surfing habits, especially when looking for educational material.

7.3. DiSC as a Learning Tool

Our survey and interview data suggest that DiSC's potential benefit would be highest in educational environments. 75% of our survey respondents used the Internet for research and 46% specifically for education and training, suggesting wide interest in these areas. Since other popular categories like email and chat are not well served by DiSC due to the dynamic nature of the content, education seems like a natural application for DiSC. Since educational users are likely to have common interests (e.g. mathematics or history) and educational information can often be relatively static (without losing its value), DiSC could be very beneficial to under-funded Ghanaian schools seeking to connect to the Internet. DiSC would also allow searching when the network is down. Our interviews also identified a pressing need in educational institutions for low cost technological solutions to mitigate resource shortages and enhance learning. Thus, our qualitative data, coupled with our technical data from BusyInternet, supports our existing hypothesis that DiSC will be useful in educational institutions. Further research would be required to identify specific requirements of school computing labs and understand the relevant contexts where DiSC could provide the most value.

7.4. Adoption Potential of DiSC

We now discuss the potential for adoption of DiSC in Ghana, given existing conditions. Our method is to analyze the political, economic, societal, and technological drivers of technology adoption in Ghana and evaluate DiSC's potential for widespread adoption, particularly in educational institutions. For this analysis we have relied primarily on interviews with experts in Ghana and outside research.

Political Trends

The general political environment in Ghana is stable, as evidenced by a recent peaceful transfer of power and free elections, with very little civil strife. In terms of technological development and policies, our interviews with government officials indicated a genuine commitment to technological progress, particularly in promoting full connectivity and access.

Since much of the telecommunications infrastructure and equipment is or was formerly owned by the state, the critical role of government in encouraging wider adoption of technology cannot be overstated. Besides having access to large amounts of capital, the government is uniquely positioned to solve the inevitable coordination problems that arise when new technologies are adopted.

The Ministry of Communications and Technology, led by Mr. Kan Dappah, was established in 2001, to administer and develop ICT related policies. Their signature policy initiative is to establish full connectivity in Ghana by 2006. [<http://www.iconnectghana.org/pages/ictinghana/index.asp>] [<http://www.ghana.gov.gh/pbcopin/index.php>] While these goals seem somewhat ambitious, the general policy direction favors low cost innovations like DiSC, especially where international aid can be used to supplement the adoption of DiSC in schools, including those in rural or underserved areas.

On the other hand, many of our experts complained about excessive government interference in business, citing a recent example of Voice Over IP (VoIP). While many entrepreneurs have sought to establish VoIP services, the government is reluctant to give up its stranglehold on long distance telephony, from which it extracts huge tariffs. Entrepreneurs argue that many people are using VoIP illegally anyway, and if the government recognized it and regulated it, they could recover much of their lost revenue. Needless to say, the government has been slow to recognize this. Thus, progress on monetizing VoIP has been slowed by the government.

In summary, while the government often provides onerous obstacles for the development of private enterprise in Ghana, an education-focused solution like DiSC would largely rely on and benefit from significant government involvement. During our conversation, we briefed Minister Dappah and his staff on DiSC. The officials quickly and clearly saw its benefits and would likely be helpful if DiSC could be implemented during their tenure. Since

DiSC would not be implemented in the classic private enterprise model, government cooperation may be more useful than usual.

Economic Trends

The economic drivers of technology adoption in Ghana are complicated and difficult to evaluate. A theme that reoccurred several times during our interviews was the lack of available capital to fund innovative technological ideas (e.g. a venture capital market). Since private firms often promote new technologies and promote adoption in the developed world, some believed that technology adoption in Ghana would be hampered by the lack of innovative entrepreneurial companies in the ICT sector. Large companies are more apt to focus on existing ideas and expanding their service, rather than developing new technologies. One reason for this is that in many developing nations, the large, and formerly state owned enterprises that run telecommunications are insulated from competition due their size and government patronage, reducing the need to innovate.

Obtaining loans is also problematic because credit history is difficult to collect and retrieve, making simple exercises like paying a security deposit for a new apartment very difficult. While various arrangements can be made to mitigate these issues, the fact remains that Ghana's financial markets are underdeveloped, resulting in small amounts of risk capital for developing new ideas. One positive development has been an inflow of capital from foreign investors, like Mark Davies of BusyInternet, who also incubates young companies with new ideas.

When we visited an Internet Cafe near the University of Ghana, we found it completely empty, save for the owner, who explained to us that the Internet was down, and that we could come back later to use the computers. We explained to him how DiSC would enable offline searching, since frequently viewed websites would be cached. He immediately recognized the value of the system and asked us to provide him more information.

By offering DiSC for offline searching, bandwidth demand decreases and Cafe owners and university computer labs could charge a reduced price for DiSC searching. This model would allow Cafes and computing labs to stay open during predictable hours and perhaps lead to cost reductions, a worthy goal, since Internet access is still quite expensive for the average Ghanaian. Our research demonstrates that this economic model is suitable for DiSC and for Ghana, and further implementation will hopefully yield further insight.

DiSC is also economically feasible because of the underlying open source software and because it takes advantage of existing hardware. Additionally, since many international institutions are targeting low cost technological solutions with applications for education or government, the cost of administering DiSC could potentially be partially offset by international grants.

Societal Trends

Societal and cultural norms among Ghanaians have a significant impact on their individual decision to adopt information- and communication-based technologies. Our research suggests that Ghanaian culture and traditions have enabled the adoption of new technologies at the individual level, motivated in large part by the desire to communicate with family, but also by the widely held belief that Ghana must lead Africa in economic and technological development. In addition, these norms are also manifest at the institutional level as evidenced by the development of dedicated IT research centers, such as the Kofi Annan Center for IT Excellence, and new private universities, which are increasingly focused on developing graduates with strong critical thinking, information technology and computer programming skills. Ghanaians are eager to compare themselves to India, China, and other emerging nations, rather than their African peers, especially when it comes to Internet adoption and related technology.

Our research suggests that Ghanaians have a very positive view towards the Internet and related web applications. They view Internet technology as a method of sustaining strong relationships among their friends and members of their extended families. This was especially

true for Ghanaians who have family members or friends living abroad. Based on our survey results, users primarily use the Internet as a medium for communication as evidenced by their predominant usage of email and chat services. 94% and 54% of our survey respondents used the Internet to email or chat, respectively. While communication technologies were the most widely adopted, Ghanaians are expanding their use of the Internet as a resource to enhance their own education and keep abreast of national and international current events.

Beyond our anecdotal research, there exist several data points that support the notion that Ghanaians embrace technology. First, cell phone penetration among Ghanaians is approximately 33%. While much of this penetration is concentrated in urban centers, adoption is permeating rapidly to rural areas given the lower deployment costs of wireless communication infrastructure. Through our observations, focus groups, and surveys, we discovered that Ghanaians might more quickly understand and utilize new technologies than our preconceived notions would have indicated. Cost is the major driver of adoption, and low-cost solutions are critical.

Second, Internet Cafes have taken Ghana by storm in mostly urban but also rural settings. As of 2003, there were 1000 Internet Cafes located across Ghana. [www.icconnectghana.org] This is expected to increase significantly over the next few years as both the demand for Internet Cafes rise, and the cost of both computing and Internet access decrease. Surely the Cafes are having a noticeable impact on Ghanaians, especially among youth, who visit the Cafes at regular times throughout the week to connect with friends, listen to music, or do research.

While the general population of Ghana is increasingly becoming pro-technology, the educational infrastructure required to support such a paradigm shift is lagging. Based on our interviews with academic leaders at both the University of Ghana, the largest public university in Ghana, and Ashesi University, a recently launched private university, we learned that the traditional culture of learning emphasizes rote memorization as opposed to more useful problem

solving skills. In addition, resources necessary to train a technologically skilled workforce are lacking. For example, at the University of Ghana, the computer science computing lab is not networked and maintains computers that are at least ten years old. Additionally, students must handwrite their computer programming projects rather than programming on a computer due to lack of computing resources at the University Ghana. Finally, students have to pay to access the Internet on campus.

In contrast to the University of Ghana, Ashesi University focuses on teaching students critical thinking and problem-solving skills. At Ashesi, students complete a four-year program that includes a liberal arts curriculum with emphasis on reading and writing skills, a rigorous analytical core of mathematics and computer programming classes as well as business and entrepreneurship classes. Students have a choice of whether to specialize in computer science or business, but each student graduates with the core skills and rigorous academic background to become leaders in society. Some other private universities near Accra and Kumasi are adopting similar models to address the need to produce graduates who can think critically and solve complex business and technical problems. However, these universities are still few, while the University of Ghana, the nation's largest public university system, still enrolls the majority of students. In terms of high schools, our interviews revealed a severe shortage in funds for basic educational needs, much less new computer hardware and software at the secondary level and below.

Technological Trends

Rapid technological changes related to Internet Access are taking place in Ghana that will influence the potential adoption of DiSC. Some positive developments bode well for the enabling of widespread Internet access within the country. However, other changes may impede the spread of Internet access and related services.

The majority of Internet Service Providers in Ghana use a satellite connection (VSAT) to connect to the Internet backbone. VSAT is a geo-synchronous orbit, high latency satellite

technology that enables an organization to connect "directly" to the Internet. There are no land links for ISPs in Ghana. A typical transmission from Ghana to the US over the Internet consists of the following hops:

1. *Ghana to Portugal (via satellite)*
2. *Portugal to the United Kingdom (via fiber)*
3. *United Kingdom to the United States (via fiber)*

While satellite connections provide Ghanaian ISPs with high-speed access to the Internet backbone, the tradeoffs in terms of cost and reliability are significant. The most expensive component of the connection is the first segment: Ghana to Portugal. This is largely because of the high fixed costs associated with deploying satellites. However, without other alternatives, ISPs must pass this cost to their customers, namely local businesses and educational institutions. The cost of Internet access will not come down in the near future until either more competitive solution enter the marketplace or the government intervenes by providing subsidies.

The high cost of Internet access has several implications for business owners, Internet Cafe operators, university network administrators, and end users. Business owners are less likely to use the Internet and, as a result, become less competitive. Similarly, Internet Cafe operators must achieve a certain degree of scale to have the ability to offer high-speed connectivity. This usually translates in poor long-term financial sustainability for Cafes. While many Cafes have recently opened in both urban and rural areas, it is more than likely many will not be able to survive unless the cost of Internet access decreases or they are able to diversify into other lines of business. Lastly, universities such as the University of Ghana continue to charge their students a fee to use the Internet to subsidize the costs of providing students high-speed connections. This places an unnecessary economic strain on students, many who already are struggling to fund their education.

In addition to deterring widespread adoption of the Internet, the high cost of Internet access produces impediments to the development of an underlying local web content and services industry. This is augmented by the fact that ISPs have little incentive to cooperate. First, there is neither a national Internet backbone nor local Internet exchange points that allow for a more cost-effective method of sharing and transmitting local web pages and data. Currently, ISPs who chose to share local data must establish wireless point-to-point solutions. As a result, this often leads to a negotiation process among ISPs that is mired by personal political and economic agendas. Hence, it is difficult to scale peer-to-peer connections. A better solution would be to create local Internet exchange points in major urban centers. The exchange point would need to act independently, charging ISPs a pre-determined locale-specific market rate for exchanging information. Rather than having ISPs connect via satellite to exchange local content and information, ISPs could achieve this at much more cost efficient levels locally. Hence, establishing a local Internet exchange point would significantly lower the entry costs for local entrepreneurs who are interested in creating local content and services.

Because of these obstacles, a technology such as DiSC could potentially alleviate the significant upfront and ongoing economic costs of providing high-speed Internet access. DiSC provides network administrators the opportunity to displace some queries to the cache and as a result open up bandwidth for 'new' non-cached searches and content. Additionally, DiSC could reduce the overall demand for bandwidth by an organization during peak hours. Last but not least, DiSC enables static information that currently sits on servers in the US or Europe to be much closer to the end user, which translates in more rapid document recovery and improved user satisfaction. Lastly, DiSC also provides an alternative for providing Internet access in rural settings. A modification to DiSC could allow rural communities to have access to a 'cached' and relevant subsection of the Internet that is updated via narrowband (i.e. UHF) technologies every night.

Thus, the current technological state of Internet access seems to favor the adoption of DiSC. While the Ghanaian government and the private sector are in the process of deploying widespread Internet across the country, the opportunity for DiSC to reduce costs, improve end-user satisfaction, and spur the development of local content sites still remains large.

After analyzing these 4 key areas, we find the adoption potential for DiSC to be promising, with some considerable obstacles. Only future research will determine whether these barriers can be overcome.

8. Conclusion

We began our research with a clearly defined technical problem, poor Internet connectivity in developing nations, and a candidate technology, DiSC. DiSC's ability to cache frequently viewed documents and improve the speed of document retrieval was hypothesized to have benefits in a developing nation computing environment characterized by intermittent connections. Our goal was to submit DiSC to testing in the field and survey users on their experiences using our software. Since we also had the wider implementation of DiSC and other technologies in mind, we conducted several interviews with thought leaders to better understand the broader economic, political, and social trends influencing technology adoption.

The technical data we collected is revealing. First, even environments with relatively expensive satellite connections to the Internet exhibited intermittent and non-uniform behavior. This suggests that DiSC's ability to mask network outages by facilitating offline searching and browsing will continue to be beneficial even as more modern networks are deployed. Since DiSC searches the local cache, it represents a marked improvement over existing caching systems, where the information is stored but not easily accessible.

Next, our trace data indicates that in the context of one Internet Cafe, there is significant locality of reference for document retrieval, but not for searching. We believe there will be

greater locality of reference for both search and document retrieval in educational environments, and hence that DiSC might still be quite useful in many educational environments.

We also discovered that an intuitive interface would be critical for future versions of DiSC and other technologies. Network administrators have enough problems on their hands without another headache from difficult to configure software.

The survey results were also illuminating. The most interesting finding was that many respondents used the Internet for education and training, which is a fact that bodes well for the utility of DiSC. Since DiSC works the best in computing environments characterized by a homogenous set of users and by static information, the education and training space is a promising potential target for a wider roll out of the software (although we were unable to confirm that indeed education and training users actually utilize more static data than other types of users, due to the strike mentioned above).

Interviews were also important in that they gave us a unique perspective on technology in Ghana. While the DiSC technology was the fundamental focus of our research, the interviews elucidated an unavoidable fact: economic, political and social development impact technology adoption as much as the underlying technology itself. While we discovered a considerable interest in DiSC, wide adoption of a searchable cache technology would require coordination and compromise with many different partners in Ghana and possibly Africa as a whole.

The broader technological trends seems to suggest the need to push more local content to “the edge” of the network by creating incentives to develop Ghanaian content and facilitate Internet exchange points. DiSC is just one mechanism to bring information closer to the end user, but it has tremendous potential to mitigate concerns over intermittent connectivity, high costs, and slow connection speed. In future research, we plan to investigate other techniques for pushing data to edge of the network, as well as gather more extensive survey and technical data on DiSC.

While many observers of Africa stress the need for Africans to “catch up” with the developed world, we view the further development of software for developing countries in the opposite way: software needs to “catch up” to the realities of modern Ghana. We hope that DiSC will be a strong step in this direction.

9. References

- Bakr, Omar, and Matulya Bansal, Fred Cheng, and R.J. Honicky. 2003. “*A Distributed, Searchable Cache for Tightly Coupled Networks*”, *Working Paper*.
[http://www.soe.ucsc.edu/~honicky/disc_paper.pdf]
- Foster, William and Seymour Goodman, Eric Osiakwan, Adam Bernstein. 2003. “*The Internet in Ghana*”, *Working Paper*
- Lucas, Henry C. Jr. and Richard Sylla. 2002. “*The global impact of the Internet: widening the economic gap between wealthy and poor nations?*”, *Working Paper*.
- National Communications Authority of Ghana, 2000
- [<http://www.ghanaweb.com/GhanaHomePage/it/>]
- [<http://www.iicd.org>]
- [<http://www.cia.gov/cia/publications/factbook/geos/gh.html>]
- [www.iconnectghana.org]
- [<http://www.iconnectghana.org/pages/ictinghana/index.asp>]
- [<http://www.ghana.gov.gh/pbcopin/index.php>]

10. Appendix

10.1. Respondent Survey Data

10.1.1. Survey by Location

| Count of Number on Survey | | |
|---------------------------|-------|------------|
| Name of Internet Cafe | Total | % of Total |
| Arms Best | 11 | 5.7% |
| BusyInternet | 103 | 53.4% |
| Hotcom | 4 | 2.1% |
| Net Connect | 43 | 22.3% |
| Queens | 27 | 14.0% |
| Royal | 5 | 2.6% |
| Grand Total | 193 | 100.0% |

10.1.2. Gender

| Count of Gender | | |
|-----------------|-------|------------|
| Gender | Total | % of Total |
| Male | 140 | 73% |
| Female | 51 | 27% |
| Grand Total | 191 | 100% |

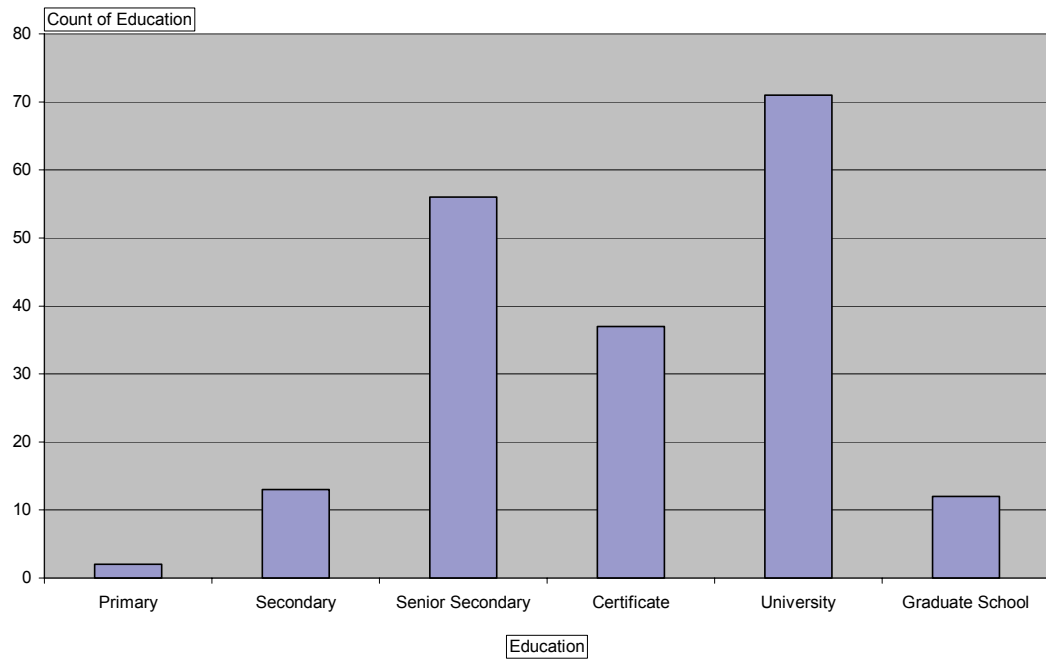
10.1.3. Student Status of Respondents

| Count of Student | | |
|------------------|-------|------------|
| Student | Total | % of Total |
| Yes | 137 | 71% |
| No | 56 | 29% |
| Grand Total | 193 | 100% |

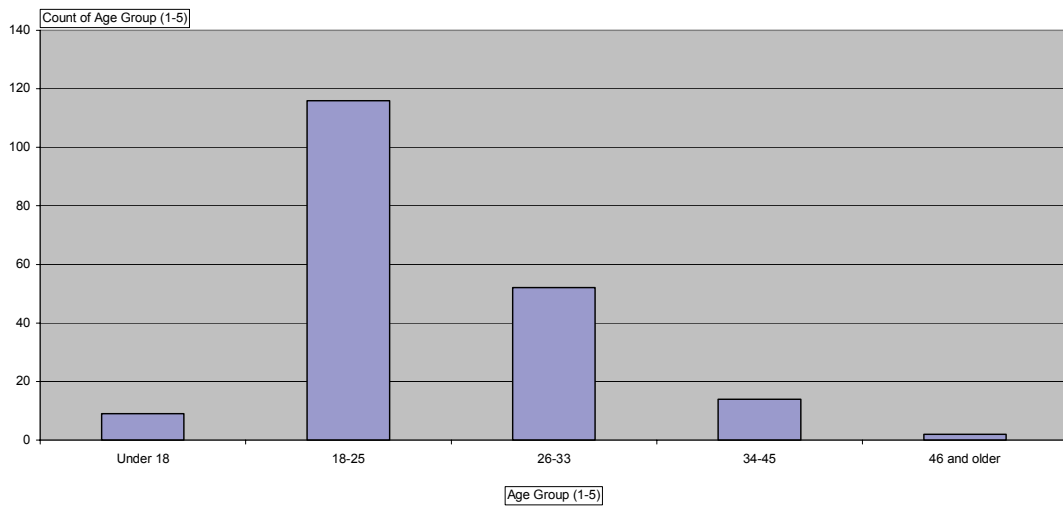
10.1.4. Visit Status of Respondents

| Count of First Visit | | |
|----------------------|-------|------------|
| First Visit | Total | % of Total |
| Yes | 16 | 8% |
| No | 177 | 92% |
| Grand Total | 193 | 100% |

10.1.5. Respondents by Education Level



10.1.6. Respondents by Age Group



10.1.7. Age by Gender

| Count of Gender | Gender | | |
|-----------------|-----------------|------|--------|
| | Age Group (1-5) | Male | Female |
| Under 18 | 4 | 5 | 9 |
| 18-25 | 80 | 35 | 115 |
| 26-33 | 45 | 7 | 52 |
| 34-45 | 10 | 3 | 13 |
| 46 and older | 1 | 1 | 2 |
| Grand Total | 140 | 51 | 191 |

10.1.8. Education by Gender

| Count of Education | Gender | | |
|--------------------|-----------|------|--------|
| | Education | Male | Female |
| Primary | 0 | 2 | 2 |
| Secondary | 9 | 4 | 13 |
| Senior Secondary | 41 | 15 | 56 |
| Certificate | 33 | 4 | 37 |
| University | 47 | 23 | 70 |
| Graduate School | 9 | 2 | 11 |
| Grand Total | 139 | 50 | 189 |

10.1.9. Computer Ownership

| Count of Own Computer | Total | % of Total |
|-----------------------|-------|------------|
| Own Computer | | |
| Yes | 143 | 74% |
| No | 49 | 26% |
| Grand Total | 192 | 100% |

10.1.10. Home Internet Access

| Count of Home Internet Access | Total | % of Total |
|-------------------------------|-------|------------|
| Home Internet Access | | |
| Yes | 40 | 21% |
| No | 151 | 79% |
| Grand Total | 191 | 100% |

10.1.11. Frequency of Internet Access

| Count of Frequency of Internet Access | Total | % of Total |
|---------------------------------------|-------|------------|
| Frequency of Internet Access | | |
| More than 3x per week | 98 | 52% |
| Between 1-3x per week | 51 | 27% |
| 1x per week | 28 | 15% |
| 1x per month | 11 | 6% |
| Less than 1x per month | 2 | 1% |
| Grand Total | 190 | 100% |

10.1.12. Online Activities by Choice

| Activity | Yes | No | % Yes | % No |
|-----------------------|-----|-----|-------|------|
| Email | 179 | 10 | 95% | 5% |
| Research | 140 | 46 | 75% | 25% |
| International News | 108 | 74 | 59% | 41% |
| Chat | 99 | 86 | 54% | 46% |
| Education/Training | 83 | 98 | 46% | 54% |
| Local News | 75 | 103 | 42% | 58% |
| Sports | 71 | 109 | 39% | 61% |
| Job/Career Search | 62 | 119 | 34% | 66% |
| Online Music/Video | 59 | 121 | 33% | 67% |
| Health Information | 40 | 136 | 23% | 77% |
| Online Games | 33 | 143 | 19% | 81% |
| Online Dating/PenPal | 30 | 146 | 17% | 83% |
| Purchase Goods Online | 22 | 155 | 12% | 88% |
| Other | 5 | 141 | 3% | 97% |

10.1.13. Email Account Ownership and Email Account Age

| Have Email Account | | Age of Email Account (Years) | |
|-----------------------------|-------|------------------------------|------|
| Count of Have Email Account | | Min | 0.10 |
| Have Email Account | Total | Max | 8.00 |
| Yes | 79 | StdDev | 1.93 |
| No | 10 | Average | 3.68 |
| Grand Total | 89 | Median | 4.00 |

89%
11%

10.2. Interview List

We interviewed the following individuals from various organizations to gain a better understanding of the political, economic, societal, and technological forces that are affecting technology adoption in Ghana.

| Name | Organization | Title |
|-----------------------|--|------------------------|
| Private Sector | | |
| Kwaku Boadu | Arrow Networks | CEO |
| Estelle Akofio-Sowah | BusyInternet | Managing Director |
| Mark Davies | BusyInternet | Founder/CEO |
| Sabra | Open Source Evangelist | Independent Consultant |
| Journalists | | |
| Nick Thompson | BBC News, Africa | Journalist |
| Gregg Zachary | Red Herring, Wall Street Journal | Editorial Director |
| ERIC m.k. Osiakawan | Reuters Fellow, Stanford University | ICT Consultant |
| Government | | |
| David Gyewu | Ministry of Communications and Technology, Republic of Ghana | Deputy Minister |
| Dorothy Gordon | Kofi Annan Center for IT Excellence | Director General |
| Education | | |

Ayorkor Mills-Tettey

Ashesi University

Mildred Wulff

Ashesi University

Nana Apt

Ashesi University

Anaba Alemna

University of Ghana, Legon

Ms. Dodoo

University of Ghana, Legon

Professor Jacob Songsore

University of Ghana, Legon

Visiting Professor,

Computer Science

Director of Business

Operations

Dean of Academic Affairs

Head Librarian

Deputy Librarian

Dean of Graduate

Studies

10.3. *Survey Instrument*

We conducted 192 surveys using the following survey. Respondents were asked to complete the survey upon their completion of their Internet Cafe session.

Date: _____

Time: _____

INTERNET USER SURVEY

1. What is your country of origin?

| | |
|---------------|--|
| ANSWER | |
|---------------|--|

2. What town do you live in?

| | |
|---------------|--|
| ANSWER | |
|---------------|--|

3. Was this your first visit to the CyberCafe?

| | |
|------------|--|
| Yes | |
| No | |

4. Are you a student?

| | |
|------------|--|
| Yes | |
| No | |

5. If you are a student, what subject do you study?

| | |
|---------------|--|
| ANSWER | |
|---------------|--|

6. If you are not a student, what is your current or most recent occupation?

| | |
|---------------|--|
| ANSWER | |
|---------------|--|

6. What is your gender?

| | |
|---------------|--|
| Male | |
| Female | |

7. What age group are you in?

| | |
|---------------------|--|
| Under 18 | |
| 18-25 | |
| 26-33 | |
| 34-45 | |
| 46 and above | |

8. Please select the highest level of education that you have completed?

| | |
|--------------------------------|--|
| <i>Primary School</i> | |
| <i>Secondary</i> | |
| <i>Senior Secondary School</i> | |
| <i>Certificate</i> | |
| <i>University</i> | |
| <i>Graduate School</i> | |

9. Do either you or your family own a computer?

| | |
|------------|--|
| <i>Yes</i> | |
| <i>No</i> | |

10. Do you have access to the Internet from home?

| | |
|------------|--|
| <i>Yes</i> | |
| <i>No</i> | |

11. If you have Internet Access at home, what type of connection is it?

| | |
|---|--|
| <i>Dial-Up</i> | |
| <i>DSL/Cable</i> | |
| <i>Other (Please specify: _____)</i> | |

12. How often do you use the Internet to access websites?

| | |
|-----------------------------------|--|
| <i>More than 3 times per week</i> | |
| <i>Between 1-3 times per week</i> | |
| <i>Once per week</i> | |
| <i>Once per month</i> | |
| <i>Never</i> | |

13. On average, how much time do you spend on the Internet per visit?

| | |
|-----------------------------|--|
| <i>30 minutes or less</i> | |
| <i>30 minutes to 1 hour</i> | |
| <i>1 – 2 hours</i> | |
| <i>2 – 3 hours</i> | |
| <i>3 hours or more</i> | |

14. Please name five unique websites that you visit most frequently? Why do you visit each of these sites?

| Website | Reasons Why You Visit Site |
|---------|----------------------------|
| | |
| | |
| | |
| | |
| | |

15. Please name a few Ghanaian websites that you visit?

| Website | Reasons Why You Visit Site |
|---------|----------------------------|
| | |
| | |
| | |

16. How often do you use THIS CyberCafe?

| | |
|-----------------------------------|--|
| <i>More than 3 times per week</i> | |
| <i>Between 1-3 times per week</i> | |
| <i>Once per week</i> | |
| <i>Once per month</i> | |
| <i>Never</i> | |

17. What activities do you usually use the computer and Internet for?

Please select all that apply.

| | |
|------------------------------|--|
| <i>Email</i> | |
| <i>Chat</i> | |
| <i>Research</i> | |
| <i>Online Dating/Pen Pal</i> | |
| <i>Local News</i> | |
| <i>International News</i> | |
| <i>Sports</i> | |
| <i>Education/Training</i> | |
| <i>Purchase Goods Online</i> | |
| <i>Online Music/Video</i> | |
| <i>Online Games</i> | |

| | |
|---|--|
| <i>Email</i> | |
| <i>Health Information</i> | |
| <i>Job/Career Search</i> | |
| <i>Other</i> <i>Please</i> <i>specify:_____</i> | |

18. Do you have an email account?

| | |
|------------|--|
| <i>Yes</i> | |
| <i>No</i> | |

19. How long have you had this account?

| | |
|---------------|--|
| ANSWER | |
|---------------|--|

20. On a monthly basis, how much money do you spend to access the Internet?

| | |
|---|--|
| <i>Total Monthly Amount</i> | |
| <i>Monthly Amount at THIS CyberCafe</i> | |

21. Do you operate your own business?

| | |
|------------|--|
| <i>Yes</i> | |
| <i>No</i> | |

22. If so, what is the nature of your business?

| | |
|---------------|--|
| ANSWER | |
|---------------|--|

23. Do you use the *CyberCafe* for work-related activities?

| | |
|------------|--|
| <i>Yes</i> | |
| <i>No</i> | |

24. Please rank the following attributes in terms their importance on your decision to use the services at this *CyberCafe*?

| <i>Attribute</i> | Very Important | Somewhat Important | Not Important |
|--|----------------|--------------------|---------------|
| <i>Fast Internet Connection</i> | | | |
| <i>Cost of Internet Access</i> | | | |
| <i>Customer Service</i> | | | |
| <i>Networking with business people</i> | | | |
| <i>Location</i> | | | |

| Attribute | Very Important | Somewhat Important | Not Important |
|--|-----------------------|---------------------------|----------------------|
| <i>Diversity of related services</i> | | | |
| <i>Reputation/Brand</i> | | | |
| <i>Short Waiting Time for a computer</i> | | | |
| <i>Other Please specify:</i> | | | |