Designing two-list group program driven algorithm for channel navigation in internet protocol television

Timothy T. Adeliyi and Oludayo O. Olugbara*

ICT and Society Research Group,
Durban University of Technology,
Durban, 4001, South Africa
Email: tim.adeliyi@gmail.com
Email: oludayoo@dut.ac.za
*Corresponding author

Abstract: The internet protocol television took the advantage of the convergence of internet services to provide seamless interactivity, time shifting, video on demand and pay per view services. However, zapping delay is a critical problem that deters the switching intention of terrestrial subscribers and widespread of internet protocol television services. This article presents the design of an effective two-list group program driven algorithm that minimises the channel seek distance and channel seek time in order to reduce zapping delay in internet protocol television. The algorithm groups similar channel programs in related service categories such as news, movies, sports, music and documentary to make it easier for channels with similarly aired programs to be identified rapidly among several channels that an internet protocol television offers. The simulated results of the proposed algorithm show that the average seek distance is significantly reduced by 58% when benchmarked with the state of the art numerical ordering, frequency circular ordering and frequency interleaved ordering.

Keywords: channel navigation; internet protocol television; program-driven; zapping delay.

Reference to this paper should be made as follows: Adeliyi, T.T. and Olugbara, O.O. (xxxx) ‘Designing two-list group program driven algorithm for channel navigation in internet protocol television’, Int. J. Advanced Media and Communication, Vol. x, No. x, pp.xxx–xxx.

Biographical notes: Timothy T. Adeliyi is a PhD student at Durban University of Technology. He received MSc in Data Networks & Security at Birmingham City University, UK, 2012 and BSc in Information Technology from Crawford University, Nigeria, 2009. His research inclinations are multimedia systems, image processing and converged network

Oludayo O. Olugbara holds a first class Bachelor of Science (Hons.) in Mathematics and a PhD in Computer Science. He is currently a Professor of Computer Science and Information Technology and Executive Dean of the Faculty of Accounting and Informatics at Durban University of Technology in South Africa. He has developed a special inclination towards scientific computational methods in specialised areas of machine intelligence, mobile computing, image processing and smart city technology. He is deeply committed to the use of scientific methods to create relevant, insightful and effective solutions to critical problems facing the society.
1 Introduction

Television is one of the largest telecommunication media in the world as of today. Its signal distribution is going through rapid changes with the recent advancement of digital technology. However, the days where the television signal is delivered through radio frequency and terrestrial broadcasting are winding up to pave the way for the Internet Protocol Television (IPTV). This emerging model of television uses the internet technology to deliver TV programs on demand and it is the future of television broadcasting. The IPTV provides multimedia services such as TV programs, video, audio, text, graphics and data delivered over the Internet Protocol (IP) based networks. The networks are managed to provide the required level of quality of service, quality of experience, security, interactivity and reliability (Beyragh and Rahbar, 2016; Lee et al., 2015).

The IPTV has been explained to be a set of multimedia services that are distributed throughout the IP network, where a subscriber receives video streams through a Set-Top-Box (STB) connected to a broadband (Hamodi and Thool, 2013). Its emergence has brought about the seamless potential that has revolutionised personal entertainment and improved the way television services are aired to offer more TV contents than the satellite and terrestrial broadcasting. Due to the ubiquity of internet services, the IPTV took the advantage of IP infrastructure convergence to provide digital TV broadcasting. It provides a platform that offers the total digital television based on a service delivery model, which differentiates it from the traditional terrestrial TV services (Jimenez et al., 2015). The primary difference between IPTV and internet TV lies in the way contents are delivered. On one hand, IPTV uses a managed network for the delivery of video services (Strangelo, 2015). On the other hand, Internet TV is a television broadcast service distributed over the internet, typically to personal computers via streaming or downloadable video contents through a web browser (Ahmad and Begen, 2009). Moreover, the content aired on Internet TV is either user generated such as YouTube or it is provided by content providers such as al Jazeera, France 24 or BBC among others (Xiao and Ye, 2008).

IPTV services can be transmitted using a managed or an unmanaged approach. A managed IPTV service enables the offering of video contents by a single network operator who operates IPTV businesses or has an agreement with a video content provider who seamlessly guarantees the quality of service and quality of experience (Friedrich et al., 2010). Furthermore, a managed service is the one provided by an IPTV service provider based on the standard STB and a reserved bandwidth where providers optimise their services (Moustafa, 2012). In an unmanaged IPTV deployment, a third party via the Internet Service Provider (ISP) or a telephone company without the guarantee of Quality of Service (QoS) and Quality of Experience (QoE) offers video contents. Friedrich et al. (2010) further explain the concept of unmanaged service to mean IPTV services that are delivered through a service called over the top. This is the delivery of a multimedia service without any multi-system operator such as Amazon Prime, Netflix and YouTube.

The IPTV should offer an improved QoE than the terrestrial models, regardless of the intrinsic advantages that come with it. In view of this, it is practically essential to address the zapping delay problem often experienced in the IPTV. Zapping delay is a problem inherently caused by decoding and encoding time, STB processing delay, jitter and channel changing time (Zare et al., 2016; Nikoukar et al., 2016). The problem is often
Designing two-list group program driven algorithm

heightened by constant channel navigation in initiating a change to watch a desired channel among several available channels. The zapping delay is a major impediment to seamless deployment that affects the attractiveness of IPTV and significantly deters the switching intention of current terrestrial subscribers (Ramos et al., 2011; Quintero et al., 2015).

The purpose of this study is to introduce a new approach that eliminates channel prediction and channel driven approaches (Bikfalvi et al., 2011; Das et al., 2015) in order to address the zapping delay problem effectively (Ryu et al., 2014). This new approach proposes a two-list group program driven algorithm that enables the subscribers to select the desired channel from a group of channels according to the personalised popularity of the channel. In addition, subscribers can personalise channels in regards to the order of preferences through two separate lists of hot and cold channels. The lists of hot and cold channels are subsequently updated based on how frequent a channel is watched in relation to the time spent watching the channel. This approach can be implemented in both the managed and unmanaged IPTV through a set-top-box strategy.

2 Related work

The STB stops the packet stream belonging to the current TV channel and checks if the requested TV channel is available when a subscriber initiates a channel change request through the IPTV STB remote control as shown in Figure 1. The immediate check for the availability of an IPTV channel is done only if multiple channels are sent to the STB of a subscriber. However, if the TV channel is not available at the STB, an Internet Group Management Protocol (IGMP) leaves a message that is generated by the STB for the current channel. In addition, the IGMP join message generated for the requested TV channel is forwarded from the subscriber’s gateway through the aggregation network to an IGMP proxy server. The requested channel is then forwarded to the subscriber’s STB to be displayed on the TV set. The request for a new channel causes the zapping delay problem because of the large range of channels that are available to subscribers, which can result in long channel seek distance and high channel seek time (Manikandan et al., 2014).

Figure 1 Channel change request (see online version for colours)

In recent years, a number of schemes have been proposed to reduce the effects of zapping delay in the IPTV. For instance, Lee et al. (2014) proposed a two-list scheme that
identifies hot and cold channel lists. This scheme actively maintains a list of the hottest channels to watch in the near future by a subscriber and a list of cold channels that a subscriber is not likely to watch. The segregation of channels into hot and cold lists is based on how frequently and recently a channel has been viewed by a subscriber. Their scheme fetches the channels in the ‘hot’ list, previewed in a low resolution when a subscriber changes a channel. However, the channels in the hot and cold lists are arranged in reference to the Frequency Channel Ordering (FCO). This mechanism gives a longer channel seek distance when compared to the Frequency Interleaved Ordering (FIO) that places hot channels in the centre of a circular list according to their popularity (Lee et al., 2009).

A channel reordering approach called the Frequency Interleaved Ordering (FIO) was proposed to minimise zapping delay for subscribers to view the preferred channels with no delay (Shakin et al., 2014). FIO reorders ‘n’ total number of channels based on ranking from 1 to n, placing the most popular channels in the center of a circular list. The channels with odd ranks are placed on the left of the list whilst channels with even ranks are placed on the right of the list (Chen and Chiu, 2012). Regardless of how effective this approach is, popular channels are still adjacent to the less popular channels. Moreover, there is no assurance that subscribers are going to switch between the proposed popular channels.

Ramos et al. (2010) proposed an approach called channel smurfing that concurrently sends the neighboring channels with the requested channel to the STB. Hence, if a subscriber decides to switch to any of these neighboring channels, the switching delay is virtually zero and quality of experience is less affected. However, the challenge is that if a subscriber does not switch to a neighboring channel, zapping delay problem absolutely remains. Lee et al. (2010) improved the channel smurfing scheme by adding a predictive tuning algorithm based on the preferred channel selection behavior of a subscriber. The improved scheme predicts a channel that a subscriber is likely to watch rather than the adjacent channel. This intrinsically eliminates the uncertainty of sending a channel that a subscriber won’t be likely watching. The authors assumed that IPTV subscribers tend to keep changing channels via the remote controller, which informs the use of STB for channel prediction. The behavioral model of subscribers occurs in two forms. The first form is the channel switching behavior and the second is the surfing behavior that they used to predict the channel that a subscriber might request.

Oh et al. (2010) proposed a hybrid scheme that agglutinates channel prefetching and reordering schemes to address the zapping delay problem in the IPTV. Specifically, the adjacency and popularity based prefetching schemes were amalgamated with a popular channel reordering scheme to address the zapping delay problem. In spite of the amount of literature addressing zapping delay in recent years, there seems to be no lasting solution based on channel navigation, a gap that inherently serves as a strong inspiration for this study. In response to this gap, our proposed algorithm addresses the overall problem of zapping delay by developing an effective two-list group program driven algorithm to effectively reduce the channel seek distance and channel seek time.

3 Material and methods
The IPTV offers several channels, but selecting the desired one in regards to the desired program is a hard problem. This is intrinsically becoming a factor that causes zapping
designing two-list group program driven algorithm

In this paper, we present in Figure 2(a)–(c) a discussion of popular channel navigation approaches for reducing channel seek time and seek distance. In Figure 2(a), the conventional Numerical Ordering (NO) is an approach where channel popularity is likely scattered and is not aligned concretely with the subscriber preferences. This makes the total seek distance before a subscriber finds a preferred channel in most cases longer. Figure 2(b) in which channels are arranged in a circular form shows the Frequency Circular Ordering (FCO) scheme that improves the NO by aligning channels based on popular preferences of subscribers.

Figure 2 Popular channel navigation approaches for reducing channel seek distance and channel seek time. (a) Numerical ordering, (b) frequency circular ordering, (c) frequency interleaved ordering and (d) two-list scheme (see online version for colours)

However, Manjunath et al. (2013) criticised this approach because of the innate weakness that it possesses in that the least cold channel is adjacent to the most popular hot channel when navigating in a circular form. They concluded that although the chances of it happening are low, it is in contradiction to the main goal of reducing total seek distance. Figure 2(c), Frequency Interleaved Ordering (FIO) addresses the weakness of the scheme in Figure 2(b), by placing the most popular hot channel at the center of the circular
list and channels with even and odd ranks on the right and left sides respectively. This approach makes the use of up and down buttons on the remote control to quickly search for the preferred channels, hence reducing the channel seek time. Figure 2(d) shows the two-list scheme to discuss further because it provides a theoretical foundation for the approach of this study.

3.1 Two-list scheme

The two-list scheme by Lee et al. (2014) identifies hot and cold channel lists where channel preferences are determined from the previous channels watched by a subscriber. In this study, we have built on the TLS by arranging channels based on their popularity, placing the most popular hot channels in a separate list in the inner circle. The channels with even and odd ranks are respectively placed on the right and left sides based on the preferences of subscribers. In addition, cold channels are placed on a separate list in the outer circle whilst hot channels are placed in the inner circle as shown in Figure 2(d). This configuration eliminates the possibility of cold channels to be adjacent to the hot channels.

Figure 3 presents a flowchart that allows the subscribers to personalise channels into hot (channels often viewed by the subscriber) and cold (channels not often viewed by the subscriber) lists according to their preference ordering (Al-Hezmi et al., 2006; Yang et al., 2015). This strategy eliminates the previous approaches of predicting the channel that a subscriber would likely prefer to watch (Oh et al., 2010). A subscriber selects a channel to view over a period of time. The channel popularity \( P \) is determined based on the viewing rate \( R \) of a channel across all subscribers and \( H(x) \) is the time-weighted frequency of watching the channel by a subscriber. The function \( H(x) \) is determined by the contribution of each past view of a channel, while \( x \) being the elapsed time for viewing the channel (Lee et al., 2014). The channel popularity determines if a channel is to be placed in a hot or cold channel list and is expressed as follows.

\[
P = RH(x)
\]  

(1)

The popularity of a channel decreases if the channel is not regularly viewed over a period of time and it increases otherwise. Decreasing in the popularity of a channel means that the popularity value is less than a given maximum threshold (Thr_Max). The threshold value is usually set by the IPTV provider whilst channel popularity increases if its value is greater or equal to Thr_Max. If the popularity of a channel in the hot list is less than Thr_Max, the channel will be moved to the cold list. Similarly, if the popularity of a channel in the cold list is greater or equal to Thr_Max, the channel will be moved to the hot list. This process is iterated in order to update hot and cold channel lists based on the predefined maximum threshold.

The channel popularity was modelled by the Zipf distribution, which has the ability to represent the skewness of channel popularity (Oh et al., 2010). The probability \( P_i \) of the \( i \)th popular channel that will be viewed by a subscriber is determined by the following Zipf distribution function.

\[
P_i = \frac{(i-1)^{1/p}}{\sum_{i=1}^{N}(i-1)^{1/p}}
\]  

(2)
Where \( n \) is the total number of channels and \( \theta \) \((0 \leq \theta \leq 1)\) is the Zipf parameter that determines the degree of the skewness of channel popularity. All channels are equally popular whenever \( \theta \) is zero, but the popularity of a channel is increasingly skewed as the value of \( \theta \) increases. Finally, the popularity is mostly skewed whenever it becomes one.

**Figure 3** The flowchart to maintain or update channels in hot and cold lists

3.2 The two-list group program driven algorithm

The conventional channel-driven algorithms such as the numerical ordering, frequency circular ordering and frequency interleaved ordering for channel navigation still present zapping delay that limits the subscribers to view the desired channel effectively because of the high channel switching rate (Zare and Rahbar, 2016). The Two-list Group Program Driven (TGPD) algorithm reported in this paper reduces zapping delay by minimising the total seek distance and total seek time. The algorithm groups similar channel programs in related service categories such as news, movies, sports, music and documentary. This strategy makes it easier for channels with similarly aired programs to be identified rapidly among several channels that an IPTV offers.

The setup phase of the TGPD algorithm allows a subscriber to choose a channel from a hot list or cold list. The algorithm shown in Figure 3 maintains the channels in the hot and cold lists based on the viewing rate and time spent by a subscriber during the watching of a program. The TGPD algorithm makes it possible that when a subscriber
chooses a channel from a program category, the channels with similar programs will be available in the group. This strategy reduces channel navigation and streamlines the channels based on the desired channel that a subscriber has requested.

**Figure 4** Prefetching process of TGPD when a subscriber chooses a channel (see online version for colours)

The TGPD algorithm selects the pre-joined channels based on their popularity. In Figure 4, a random channel switching sequence for four program channel groups is used to calculate the number of channel switches and average total seek distance in comparison with other conventional schemes. The groups are ordered using the two-list scheme and we have assumed that subscribers of popular groups may want to watch news, movies, sport and documentary. The seek distance in moving from one group to another is one and the seek distance in moving from hot channel list to the cold channel list is also one. Figure 5 shows the prefetched popular channels that interest a subscriber among the four groups of program channels.
Figure 5  Prefetched channels into hot and cold lists from the four different program groups (see online version for colours)

4 Experimental result

The simulation experiment of this study was performed by using a pseudo-randomised grouping (Manikandan et al., 2014; Lee et al., 2014; Zare et al., 2016) to validate the performance of the TGPD algorithm. We have assumed in this study that channel popularity is based on how frequent a channel is being viewed and bandwidth is steady and sufficient. Moreover, we have categorised TV channels into four groups, which are news, movies, sports and documentary based on the similar programs the channels offer. The network parameters like end-to-end delay and jitter were not considered during the experimentation to avoid their negative effects.

To determine how channels will be maintained in the hot and cold lists, the viewing rate of a particular channel is multiplied by the total number of traces. If the viewing rate of watching a channel is 0.58, the number of times the subscriber has viewed that particular channel will be 870 out of 1500 traces. Such a channel with a threshold higher than 500 will be maintained in a hot list. Similarly, if the viewing rate is 0.16, the number of times a subscriber has viewed a particular channel is 240 out of 1500 traces and such a channel will remain in the cold list. Figure 6 shows the comparison of the TGPD algorithm with the state of the art, Number Ordering (NO), Frequency Circular Ordering (FCO) and Frequency Interleaved Ordering (FIO) for channel switches when TV channel is 20. This result indicates that the TGPD algorithm gives a significant channel switching reduction when the number of TV channels is set at 20.

Figure 7 shows the comparison of the TGPD algorithm with the conventional schemes when TV channels are 64. This result indicates that the TGPD algorithm consistently gives a significant channel switching reduction with increasing number of TV channels.
The simulation experiment was performed to validate how channels are prefetched from the hot and cold lists with synthetically generated traces. The number of channels requested in the trace is 1500 and the number of distinct channels ranges from 10 to 150. Figure 8 shows the percentage of the average seek distance reduction of the TGPD algorithm when compared with other conventional schemes as the channel popularity varies by the Zipf parameter. When the Zipf parameter $\theta$ is 0, all the schemes merge to a point, which means that all channels have an equal average seek distance. Moreover, when the Zipf parameter reaches 1, the TGPD algorithm reduces the average seek distance by 58% in comparison to the other schemes.
The experimental results of this study generally imply that the TGPD algorithm is suitable for addressing zapping delay. The proposed algorithm recorded impressive results when compared to the other schemes investigated in this study because of its ability to prefetch channels from the group of programs. The algorithm will enable subscribers to surf for their desired channel within a limited time.

5 Conclusion

This paper presents an effective two-list group program driven algorithm to address the zapping delay problem in IPTV. The algorithm has been validated through a trace-driven simulation by identifying hot and cold channel lists through different categories of TV programs put together as a group. The channels are being prefetched from a group and stored in a hot or cold channel list based on their popularity. The simulation results of this study show that the performance of the algorithm reported is impressive when compared to other conventional schemes. It can serve an important tool for reducing the zapping delay time it takes a subscriber to choose the desired channel in the IPTV. In future work, we plan to practically implement the reported algorithm in a real world managed IPTV over a converged network. In addition, the reported algorithm can be studied on different broadband line speeds to ascertain its performance.

References

Ahmad, K. and Begen, A.C. (2009) ‘IPTV and video networks in the 2015 timeframe: the evolution to medianets’, IEEE Communications Magazine, Vol. 47, No. 12, pp.68–74.

Al-Hezmi, A., Rebahi, Y., Magedanz, T. and Arbanowski, S. (2006) ‘Towards an interactive IPTV for mobile subscribers’, in Proceedings of International Conference on Digital Telecommunications (ICDT’06), pp.45–45.

Beyragh, A.A. and Rahbar, A.G. (2016) ‘A novel idea on supporting mobile IPTV services over mixed DVB-h. and 3G networks’, Multimedia Tools and Applications, Vol. 75, No. 4, pp.2091–2110.
Bikfalvi, A., Garcia-Reinoso, J., Vidal, I., Valera, F. and Azcorra, A. (2011) ‘P2P vs. IP multicast: comparing approaches to IPTV streaming based on TV channel popularity’, Computer Networks, Vol. 55, No. 6, pp.1310–1325.

Chen, Y.W. and Chiu, T.T. (2012) ‘Minimizing zapping time in IPTV based on user’s interest’, International Journal of Computer and Communication Engineering, Vol. 1, No. 2, pp.77–80.

Das, S.K., Naor, Z. and, Raj, M. (2015) ‘Popularity-based caching for IPTV services over, P2P networks’, Peer-to-Peer Networking and Applications, Vol. 10, No. 1, pp.156–169.

Friedrich, O., Thatmann, D. and Arbanowski, S. (2010) ‘An IPTV service state API for converging managed and unmanaged IPTV infrastructures’, in Proceedings of 2010 IEEE International Conference on Multimedia and Expo (ICME), pp.1493–1498.

Hamodi, J.M. and Thool, R.C. (2013) ‘Investigate the performance evaluation of IPTV over WiMAX networks’, arXiv preprint arXiv: 1302.1409.

Jiménez, J.M., Canovas, A., López, A. and Lloret, J. (2015) ‘A new algorithm to improve the QoE of IPTV service customers’, in Proceedings of 2015 IEEE International Conference on Communications (ICC), pp.6990–6995.

Lee, C.Y., Hong, C.K. and Lee, K.Y. (2010) ‘Reducing channel zapping time in IPTV based on user’s channel selection behaviors’, IEEE Transactions on Broadcasting, Vol. 56, No. 3, pp.321–330.

Lee, E.W., Jiyoung, Uran, O., Kern, K. and Hyokyung, B. (2009) ‘Popular channel concentration schemes for efficient channel navigation in internet protocol television’, IEEE Transactions on Consumer Electronics, Vol. 55, No. 4, pp.1945–1949.

Lee, E., Ku, J.Y. and Bahn, H. (2014) ‘An efficient hot channel identification scheme for IPTV channel navigation’, IEEE Transactions on Consumer Electronics, Vol. 60, No. 1, pp.124–129.

Lee, S., Park, E-A., Lee, S. and, Brown, J. (2015) ‘Determinants of IPTV diffusion’, Telematics and Informatics, Vol. 32, No. 3, pp.439–446.

Manikandan, M.S.K., Saurigresan, P. and Ramkumar, R. (2014) ‘Grouped frequency interleaved ordering with pre-fetching for efficient channel navigation in internet protocol television’, Multimedia Tools and Applications, Vol. 75, No. 2, pp.887–902.

Manjunath, L. and Mastani, S.A. (2013) ‘A novel approach for increasing channel navigation in IPTV based on user’s channel selection interests’, Int J. Eng Res Appl (IJERA), Vol. 3, No. 3, pp.1331–1336.

Moustafa, H. and Zeadal. S. (2012) Media Networks: Architectures, Applications, and Standards, CRC Press, Boca Raton.

Nikoukar, A., Hwang, I.S., Liem, A.T. and Lee, J.Y. (2016) ‘Mitigating the IPTV Zap time in enhanced EPON systems’, Journal of Optical Communications and Networking, Vol. 8, No. 6, pp.451–461.

Oh, U., Lim, S. and Bahn, H. (2010) ‘Channel reordering and prefetching schemes for efficient IPTV channel navigation’, IEEE Transactions on Consumer Electronics, Vol. 56, No. 2, pp.483–487.

Quintero, M.A.R., Zachey, B. and Raake, A. (2015) ‘Annoyance and acceptability of video service responsiveness’, in Proceedings of 2015 Seventh International Workshop on Quality of Multimedia Experience (QoMEX), pp.1–6.

Ramos, F.M., Crowcroft, J., Gibbens, R.J., Rodriguez, P. and White, I.H. (2010) ‘Channel smurfing: minimising channel switching delay in IPTV distribution networks’, in Proceedings of 2010 IEEE International Conference on Multimedia and Expo (ICME), pp.1327–1332.

Ramos, F.M., Crowcroft, J., Gibbens, R.J., Rodriguez, P. and White, I.H. (2011) ‘Reducing channel change delay in IPTV by predictive pre-joining of TV channels’, Signal Processing: Image Communication, Vol. 26, No. 7, pp.400–412.
Ryu, J-h., Lee, B., Kim, K.T. and Youn, H.Y. (2014) ‘Reduction of IPTV channel zapping time by utilizing the key input latency’, in Proceedings of 2014 IEEE 11th Consumer Communications and Networking Conference (CCNC), pp.263–268.

Shakin Banu, A., Ayswarya, K., Jenifer Esther Gethsee, T. and Arockia Julie, T. (2014) ‘Reconfigurable channel reordering to reduce latency in IPTV networks’, International Journal of Innovative Research in Science, Engineering and Technology, Vol. 3, No. 3, pp.1655–1661.

Strangelove, M. (2015) Post-TV: Piracy, Cord-Cutting, The Future of Television, University of Toronto Press, Toronto.

Xiao, Z. and Ye, F. (2008) ‘New insights on internet streaming and IPTV’, in Proceedings of the 2008 International Conference on Content-based Image and Video Retrieval, pp.645–654.

Yang, J., Park, H., Lee, G.M. and Choi, J.K. (2015) ‘A web-based IPTV content syndication system for personalized content guide’, Journal of Communications and Networks, Vol. 17, No. 1, pp.67–74.

Zare, S. and Rahbar, A.G. (2016) ‘Program-driven approach to reduce latency during surfing periods in IPTV networks’, Multimedia Tools and Applications, Vol. 75, No. 23, pp.16059–16071.

Zipf, G. (1949) ‘Human behavior the principle of least effort’, An Introduction to Human Ecology, Cambridge/Massachusetts.