Ten Blue Links on Mars

Charles L. A. Clarke, Gordon V. Cormack, Jimmy Lin, and Adam Roegiest
David R. Cheriton School of Computer Science
University of Waterloo, Ontario, Canada
claclark@gmail.com, {gvcormac, jimmylin, aroegies}@uwaterloo.ca

ABSTRACT

This paper explores a simple question: How would we provide a high-quality search experience on Mars, where the fundamental physical limit is speed-of-light propagation delays on the order of tens of minutes? On Earth, users are accustomed to nearly instantaneous responses from web services. Is it possible to overcome orders-of-magnitude longer latency to provide a tolerable user experience on Mars? In this paper, we formulate the searching from Mars problem as a tradeoff between "effort" (waiting for responses from Earth) and "data transfer" (pre-fetching or caching data on Mars). The contribution of our work is articulating this design space and presenting two case studies that explore the effectiveness of baseline techniques, using publicly available data from the TREC Total Recall and Sessions Tracks. We intend for this research problem to be aspirational as well as inspirational—even if one is not convinced by the premise of Mars colonization, there are Earth-based scenarios such as searching from rural villages in India that share similar constraints, thus making the problem worthy of exploration and attention from researchers.

1. INTRODUCTION

Search and other transactional web services strive to minimize response times in order to provide a sense of interactivity and to maintain user engagement. Regardless of how efficiently we implement these services, their response times are limited by roundtrip network latency, which in turn is limited by technical and physical factors, including the speed of light. For Earth-based users the physical limits imposed by the speed of light amount to roughly a quarter of a second of delay, in the case when a packet must bounce off a geosynchronous satellite. Consider, however, the case of future colonists on Mars, who will be between 4 and 24 light-minutes away, depending on the relative positions of the two planets [5]. This paper explores a simple question: Is it possible to engineer around physical laws to provide a tolerable search experience from Mars? Our work represents a small step towards solving the broader infrastructure problems associated with Internet and WWW[1] support beyond low Earth orbit.

While Martian colonies may be a decade or more in the future, plans are being actively developed, with the public support of luminaries such as space entrepreneur Elon Musk [17] and Edwin "Buzz" Aldrin, the second person to walk on the Moon [6]. Both Mars to Stay[2] and Mars One[3] propose permanent settlement, with colonists potentially living out the remainder of their lives on Mars. While the idea of permanent settlement may seem like science fiction to some, there are substantial cost savings from permanent colonization, as opposed to a traditional Apollo-style there-and-back-again mission, since fuel and other resources for immediate return would not be required. Permanent colonists can conduct more science, over much longer periods of time, greatly increasing the benefits accrued from the mission.

Current planning assumes that colonists will simply have to tolerate communication delays, limiting their ability to use the Internet. Mars One planners assume communication will be limited to email, video messages, and the like. For other services, they currently assume [3]:

> The astronauts can use the Internet, but can only surf "real time" on a number of websites that are downloaded from Earth on the Mars habitat webserver. Every astronaut will have access to his favorite websites that way. Visiting other websites will be a bit impractical because of the delay.

While in the short term our colonists will tolerate whatever is necessary for the success of the mission, long term separation from digital life on Earth need not be one of them. Searching, surfing, and shopping should be as easy from Mars as it is from Marseille.

The primary contribution of this work is an articulation of the design space of how we might engineer search from Mars. We model the problem as a tradeoff between "effort" (waiting for responses from Earth) and "data transfer" (pre-fetching or caching data on Mars). Such tradeoffs are similar to those present in modern computer architecture where the CPU (colonist/search engine) accesses its L1 cache (Martian cache) and, on a cache miss, fetches data from disk (Earth). Although, for example, the difference in latency between a Martian cache miss and an L1 cache miss is orders of magnitude, the relative differences are not too dissimilar.

[1]Worlds Wide Web
[2]`http://www.marstostay.com/`
[3]`http://www.mars-one.com/`

WWW 2017, April 3–7, 2017, Perth, Australia.
ACM 978-1-4503-4913-0/17/04.
http://dx.doi.org/10.1145/3038912.3052625

We flesh out our design by considering two concrete tasks using publicly available data. In the first task, we build on a previous short paper [11] and explore high-recall retrieval (such as conducting a scientific survey) using data from the TREC Total Recall Track. In the second task, we simulate interactive search sessions on Mars using data from the TREC Sessions Track. In both cases, our work examines what researchers might call "reasonable baselines". In essence, we have applied standard techniques (caching/pre-fetching) to these two tasks in an attempt to set the bar for how well simple solutions scale to Martian search problems. While other potential solutions exist, this work frames the problems in a way that allows comparison between approaches in terms of colonist effort and the amount of wasted data transferred.

We conclude with a discussion of two main architectural concerns we believe should be addressed before additional inroads can be made into solving additional Martian WWW needs (e.g., watching streaming video or making purchases): 1) implementation of a general-purpose infrastructure framework to facilitate solutions, and 2) methods for simulation under this framework, supporting both real-time experimentation and accelerated simulation, while accounting for delays, intermittent communication, and other phenomena.

2. BACKGROUND AND RELATED WORK

The problem of searching from Mars is intended to be aspirational as well as inspirational. Even if one remains unconvinced about interplanetary colonization in the short term, our work remains relevant in the same way that zombie apocalypse preparations advocated by the Centers for Disease Control are instructive.[4] Like that effort, theoretical considerations about unlikely scenarios can lead to insights with more immediate impact. In fact, search from Mars can be thought of as a specific instantiation of what Teevan et al. [26] call "slow search", which aims to relax latency requirements for a potentially higher-quality search experience. Slow search explores latencies on the order of minutes to hours, similar to Martian communication delays.

Technologies developed for search on Mars have potential applications closer to home in improving search from remote areas on Earth such as Easter Island, where only satellite Internet is available, and the Canadian Arctic, where Internet access remains prohibitively slow and expensive. Our work builds on previous efforts to enhance Internet access in developing regions such as rural India, where connectivity is poor and intermittent. Thies et al. [27] explored web search over email, an interaction model that is not unlike searching from Mars. Chen et al. [10] specifically tackled the problem of search over intermittent connections, attempting to optimize the amount of interaction that a single round of downloading can enable. Intermittent connections can be modeled as high latency, which makes the problem quite similar to ours—and indeed Chen et al. used some of the query expansion and pre-fetching techniques we explore here.

In this work, we assume that a functional interplanetary Internet already exists, and that the only problem we need to overcome is latency at the application layer. This is not an unrealistic assumption as other researchers have been exploring high-latency network links in the context of what is known as delay-tolerant networking [1] and NASA has already begun experimental deployments on the International Space Station [2]. Once again, there are many similarities between building interplanetary connectivity and enhancing connectivity in developing regions. Examples of the latter include DakNet [23], deploying wifi access points on buses to provide intermittent connectivity to users along their routes and the work of Seth et al. [24] to ferry data using mechanical backhaul (i.e., sneakernet)—which isn't very different from our proposal to put a cache of the web on a Mars-bound rocket (more details later).

Even if one accepts the premise of Mars colonization, there may remain skepticism about the importance of providing web search. While challenges such as sustaining life, finding appropriate shelter, and extracting energy are no doubt paramount, the psychological health of Martian colonists is important also. As the web has become an integral part of our daily lives, we believe that replicating the experience of the web on Mars is an integral element of maintaining psychological well-being. The HI-SEAS (Hawaii Space Exploration Analog and Simulation) missions and other previous efforts, which attempt to simulate long-duration habitation on Mars, are a recognition that keeping colonists sane is just as important as keeping them alive.

Having accepted the premise of searching from Mars, let us next flesh out some of the constraints in more detail. There exist technologies built around laser-based communication where it is possible to achieve good bandwidth between Earth and Mars. The Lunar Laser Communications Demonstration achieved a 622-Mbps downlink and a 20-Mbps uplink between the Earth and the Moon [7], so something like this to Mars is technologically feasible. More bandwidth can be achieved by building more satellites, so we can probably count on "reasonable" bandwidth between Earth and Mars. The other important factor is physical transit time from Earth to Mars on rockets, which we can use as a vehicle for physically delivering a cache of data (i.e., an interplanetary sneakernet). Missions to Mars have taken between 150 and 300 days over the past half century [4], and without getting into details about orbital mechanics (tradeoffs between transit time, fuel efficiency, and the prevalence of suitable launch windows), it suffices to say physical transport between the two planets will be on the order of months with current rocket technology. Physical transport time defines a "cache invalidation" problem—as whatever data we put on a rocket runs the risk of becoming stale before it arrives on Mars.

This work builds on two previous papers that have tackled the search from Mars problem. The first is an unrefereed magazine column that to our knowledge is the first articulation of the search from Mars problem [22]. That article articulates the vision, but lacks technical depth. The second is a short paper [11] that empirically examines the high-recall problem, which provide the basis of a more detailed study we describe in Section 4.

3. SPACETIME TRADEOFFS

Achievable response times for searching on Mars requires a tradeoff between latency and bandwidth. If the available bandwidth between Earth and Mars is very large, with few restrictions on usage, searching on Mars need be little different than searching on Earth. Mars would maintain a snapshot of the Earth-based portion of the web on local servers (initially delivered by sneakernet), continuously updating it

[4] https://www.cdc.gov/phpr/zombies.htm

with the help of Earth-based crawlers. Although this cache would still (unavoidably) be 4 to 24 minutes behind Earth, a searcher on Mars would experience no lag. Of course, if a search on Mars leads the searcher to an Earth-based transactional site, or to other dynamic content, that site will still be subject to response time delays unless it too provides accommodations for extraterrestrial users. We leave this issue for future work.

Unfortunately, maintaining a snapshot of the Earth-based web means that much of the transferred data will go unseen and unused, at least until the colony gains a sizeable population. Furthermore, although details regarding communications technology are far from finalized, we imagine that bandwidth will be limited and must be used parsimoniously. While some bandwidth might be usable for speculative pre-fetching and caching, potentially wasteful usage must be justifiable by potential benefits to the colonists.

At the other extreme in this design space, if available bandwidth between Earth and Mars is very limited, with usage restricted to the most critical purposes, we can do little to improve searching on Mars. Any kind of speculative pre-fetching or caching would waste critical resources. Under these circumstances, our colonists must tolerate the lag, along with the other restrictions of pioneer life.

Since bandwidth limitations are unknown at present, we quantify tradeoffs in terms of two measurable values, both independent of bandwidth:

1. It takes a Martian longer to perform an online task, relative to the time required on Earth. She requires this additional time either because she has to wait longer for an interaction to happen, or because she does extra work to complete her task. For example, while waiting for a search result the Martian might work on some unrelated task, or she might continue to peruse the results of a previous search while she waits for the new results to arrive.
2. We can send more data to Mars, relative to the amount of data we would send to a user's interaction device on Earth. For example, we might send extra search results, web pages, etc. that the user might never actually view. Two possible techniques are to pre-fetch results and to cache a portion of the web on Mars.

We can express the first value as an "effort ratio", E, where user effort might be measured in task completion time, or via some proxy, such as the number of web pages viewed:

$$E = \frac{\text{user effort required to complete task on Mars}}{\text{user effort required to complete task on Earth}}$$

We express the second value as a "data ratio", D, where data volume might be measured in bytes, or via some proxy, such as the number of web pages transferred:

$$D = \frac{\text{data transferred to complete task on Mars}}{\text{data transferred to complete task on Earth}}$$

For interactive web search, there is a tradeoff between D and E. If we perform no pre-fetching or caching on Mars, using the search engine exactly as on Earth, we have $D = 1$ but E is maximized. If we continuously send a full web crawl to Mars, we get $E = 1$, but D is maximized. On Earth, $E = 1$ and $D = 1$ by definition.

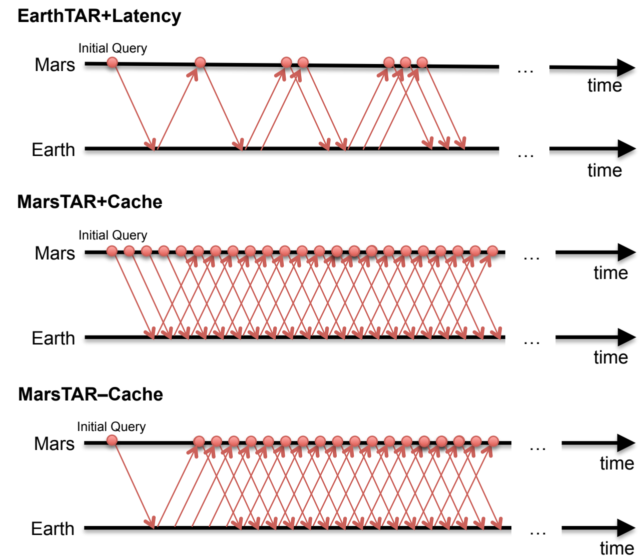

Figure 1: Illustration of various AutoTAR on Mars scenarios. Circles indicate relevance judgments.

On Mars, we trade off one against the other. While E is determined largely by the distance between the two planets and the number of roundtrip delays required to complete a task, D may be arbitrarily large, even when little interaction is required. For example, even pre-fetching pages linked from a SERP (see Section 5.2) increases the number of pages transferred by roughly a factor of ten, even if the user only clicks on a few (or no) results.

4. CASE STUDY 1: TOTAL RECALL

As an example of the tradeoffs discussed above, we extend our previous study, which considered high-recall retrieval, e.g., conducting scientific surveys, in a permanent Martian colony [11]. In this context, the Martian searcher aims to find as much relevant material as possible while minimizing the amount of non-relevant material consumed.

Our previous study examined this task using Cormack and Grossman's AutoTAR protocol [13, 15], which uses continuous active learning [12, 14] to iteratively train a classifier based upon searcher feedback collected in batches (for computational efficiency). We omit implementation details of the underlying algorithm for brevity.

In the previous study, we proposed four formulations of the AutoTAR protocol. As we rerun the experiments using more realistic measurements of time, we provide a brief description of these formulations (See Figure 1):

- **EarthTAR**: The baseline result of running AutoTAR on Earth—the upper bound of performance.
- **EarthTAR+Latency**: Results of running AutoTAR from Mars without any attempts to hide latency, i.e., the searcher waits between batches of new documents to assess.
- **MarsTAR+Cache**: Two versions of AutoTAR are running, one on Earth and one on Mars. The Martian AutoTAR begins by running on a cache that has been shipped to Mars (ahead of time). This cache is incrementally expanded as Earth identifies new potentially-relevant documents and ships them to Mars. Earth runs its own version

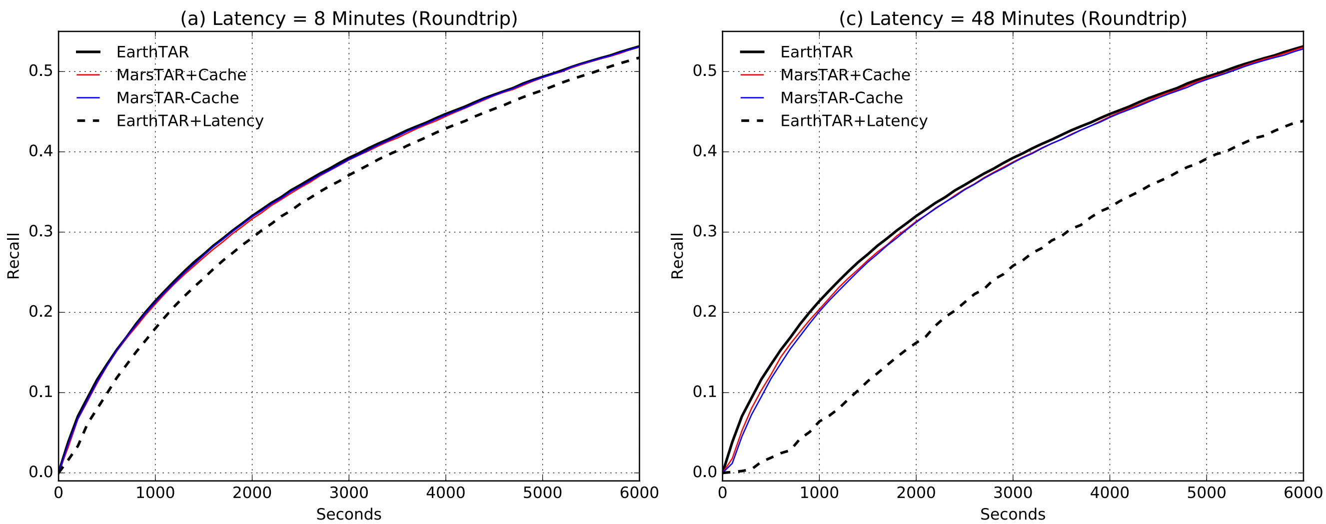

Figure 2: A comparison of the four formulations of AutoTAR with 8-minute (left) and 48-minute (right) roundtrip delays.

of AutoTAR on the entire collection, as trained by Martian assessments received after a delay.

- **MarsTAR−Cache**: As above, except there is no pre-existing Martian cache. Thus, after the initial query, the Martian must wait for a roundtrip latency before she can begin assessing documents received from Earth.

The last three formulations were compared to the EarthTAR upper bound to examine the impact of various roundtrip latencies. As before, we used the Reuters Collection Volume 1 (RCV1) [21], which comes with a fully labeled training and test set split over 103 topics. The training portion of the collection (the chronologically first ∼24,000 documents) was used as the Martian cache where applicable. Our previous study made the simplifying assumption that the time to judge a document was one light-minute, i.e., in four light-minutes of latency between Earth and Mars, the searcher can judge four documents. This assumption makes for easy simulations but is not necessarily realistic. In this paper, we adapt a simple model of reading speed for document review [25] to calculate the time needed for our Martian searcher to judge a document. More precisely, this model predicts that judging a document containing l words takes $0.018l + 7.8$ seconds.

The previous assumption that reading time is equivalent to transmission time meant that arbitrary latencies between Earth and Mars, ranging from 5 units of time to 500 units of time, could be used. In our revised experiments, we use roundtrip latencies of 8 and 48 minutes—these function as minimum and maximum possible delays based on Earth/Mars orbits.

Figure 2 plots the recall achieved as a function of the amount of time spent searching, including any perceived latency for the two levels of roundtrip latency. These results agree with those reported previously but provide a more accurate picture of the Martian user experience. It is apparent that for this task, which is tolerant of interaction delays, a small Martian cache is sufficient to achieve comparable performance with Earth. Even with no cache, Mars is quickly able to catch up to Earth-like effectiveness. Furthermore, we note that in the context of high-recall retrieval there is no wasted transmission. That is, to be sure all relevant material has been identified, a searcher must exhaustively examine the entire corpus, regardless of the underlying retrieval protocol, rather than traditional web search where a searcher may examine only one or two documents returned in a SERP. It is worth noting that at any given point in time, the MarsTAR solution may have resulted in additional non-relevant documents being received by Mars when compared to EarthTAR (i.e., $D > 1$). The plots show, however, that such a discrepancy appears to be relatively small.

5. CASE STUDY 2: SEARCH SESSIONS

In our second case study we use an existing query log to examine the impact of searching from Mars and to explore pre-fetching and caching techniques for remediating Earth-Mars latencies. For each search session in our log, we plot the number of pages transferred to the user's interaction device against the total time of the session. For searching on Earth, these numbers come directly from the log, since the sessions actually did take place on Earth. For searching on Mars, we add an appropriate delay for each interaction with the search engine. We examine the two extreme cases, an 8-minute delay when the planets are at their closest, and a 48-minute delay when the planets are farthest apart. For this simple simulation, we assume the user waits (or works on some other task) during each interaction cycle.

Academic research into web search is hampered by the paucity of query log data, particularly data for complete search sessions. To address this need, the TREC Session Track created test collections and evaluation measures to study search across multi-query sessions [19, 20, 18, 8, 9], as opposed to single-shot queries. As part of this effort, the track organizers gave specific search tasks to recruited users, recording queries, clicks, dwell times, and other information as the users conducted these search tasks. The track ran for five years (TREC 2010–2014). We used TREC 2014 data for our experiments [9].

For TREC 2014, track organizers recruited users through Amazon Mechanical Turk, recording 1,257 unique sessions, comprising 4,680 queries and 1,685 clicks. Users conducted searches with the Indri search engine over the ClueWeb12 collection, a crawl of 733 million pages from the general web

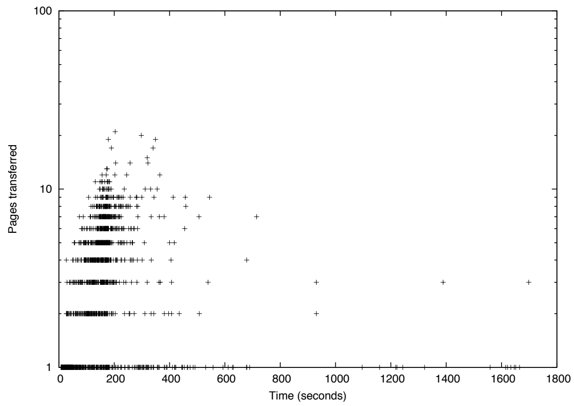

Figure 3: Search sessions on Earth: Each point indicates the pages transferred and total time for a single session.

gathered in early 2012.[5] While the size of this collection is modest by commercial standards, and the size of the log is dwarfed by a few milliseconds of commercial search, it has proven to be a valuable resource for understanding user behavior across sessions [28].

5.1 Baselines

Earth-based interactions are taken directly from the Session Track log. Figure 3 plots all sessions, with each point representing a single session. Session duration is plotted on the x axis and the number of pages transferred is plotted on the y axis. For the purposes of counting pages transferred, a SERP counts as a single page, as does a click.

Points falling along the x axis represent searches where the user issued only a single query and did not click on any results. Most sessions take under ten minutes, with a few taking nearly a half hour. In the log, session duration is assumed to start when the user begins to consider the search problem, and not when the first query is issued. We retain this approach in our experiments.

As the simplest simulation of searching from Mars, we can replay the session log, assuming the user waits (or does other work) after each query and click while the request is sent to Earth and the response is returned to Mars. Figure 4 plots sessions under minimum and maximum delay times. In these plots, delays clearly dominate interaction times, especially with a worst-case 48 minute delay. These simulations do include some very basic caching. If a query is issued multiple times or if a page is clicked multiple times, we assume the result is fetched only once.

Table 1 shows average transfers and average session duration for various scenarios, along with effort ratios (E) and data ratios (D) as defined in Section 3. Average effort ratio (E) essentially grows linearly with roundtrip time, i.e., the lag seen by the user. Data ratio (D) actually drops slightly, since we do not assume caching in an Earth-based browser. If we had, we would have $D = 1$ in all cases.

5.2 Pre-fetching

How might we begin to hide latencies associated with searching from Mars? After the initial query in a session, we might attempt to predict the user's needs and pre-fetch pages required for the remainder of the session, potentially reducing E at the cost of increasing D. We try three simple approaches: pre-fetching pages linked directly from SERPs (up to ten), pre-fetching additional related pages (perhaps several thousand) along with the pages linked from SERPs, and expanding with query suggestions and returning associated SERP pages.

Location	Lag (min)	Average time (sec)	Average pages transferred	Effort ratio (E)	Data ratio (D)
Earth	0	172.323	3.940	1.000	1.000
Mars	8	2046.118	3.904	15.334	0.995
Mars	48	11415.092	3.904	87.005	0.995

Table 1: Average performance for Earth-based and Mars-based sessions under various delay scenarios, with no pre-fetching or caching.

Location	Lag (min)	Average time (sec)	Average pages transferred	Effort ratio (E)	Data ratio (D)
Earth	0	172.323	3.940	1.000	1.000
Mars	8	1436.667	23.593	11.263	7.477
Mars	48	7758.385	23.593	62.578	7.477

Table 2: Average performance for Earth-based and Mars-based sessions under various delay scenarios, with SERP pre-fetching.

SERP Pre-fetching. As our first attempt at reducing E, we pre-fetch result pages linked from SERPs, under the assumption that the user will click on at least some of them. Indeed, pre-fetching of result pages is so obviously sensible that we cannot imagine supporting search on Mars without at least this optimization, unless bandwidth limitations are extremely severe. Here we pre-fetch only the pages directly linked from SERPs in the log, but we might imagine going further, perhaps by loading more of the linked site or by pre-fetching deeper in the results list.

Figure 5 plots individual sessions using SERP pre-fetching, and Table 2 shows average effort and data ratios. While most SERPS in our query log link to ten pages, D is less than 10 due to caching effects. E increases linearly with lag, but values are at 25% or more below those in Table 1. That is, we save around 25% effort at the cost of transferring around seven times more data than necessary.

Topical Pre-fetching. A simple way to extend SERP pre-fetching would be to go deeper in the ranked list, perhaps uploading a large set of topically-related pages in response to each query, along with the pages linked directly from the SERP. Even when a Martian user issues a further query, requiring another roundtrip delay, these pages would allow the user to further explore her topic while awaiting additional results from Earth. If large numbers of related pages are uploaded, query reformulations can also be issued against these local pages, perhaps allowing Martian users to satisfy their information needs without waiting.

To explore the potential for topical pre-fetching, we indexed the (complete) ClueWeb12 collection using the Lucene search engine. As queries appear in the log, we execute them against this index with BM25 ranking, and then assume that the top k documents are sent to Mars, along with documents linked from the SERP. Since we are using different

[5] `http://lemurproject.org/clueweb12/`

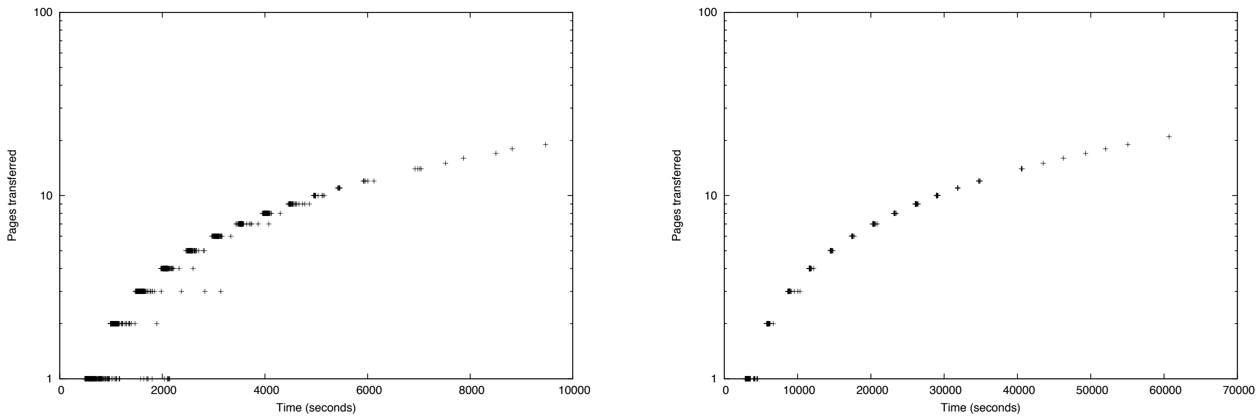

Figure 4: Search sessions on Mars with 8-minute (left) and 48-minute (right) roundtrip delay: Each point represents a single session with no pre-fetching or caching.

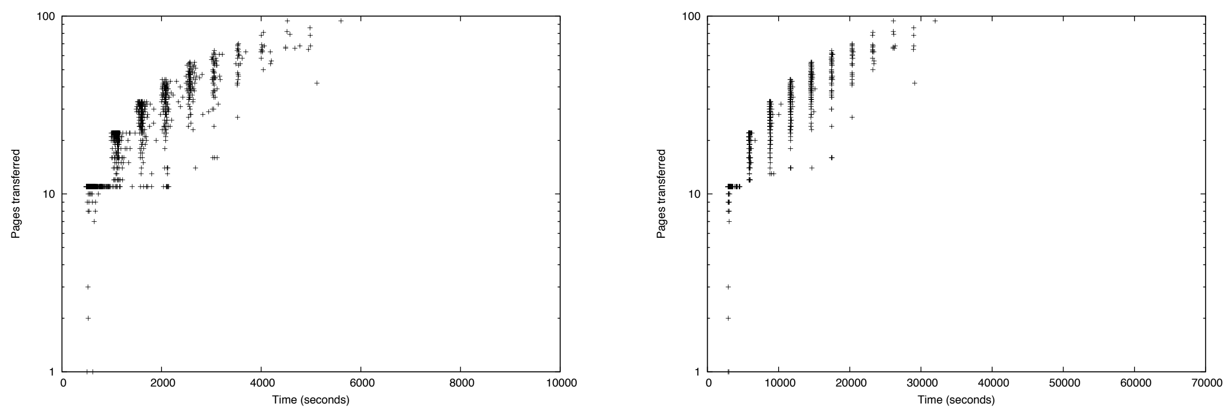

Figure 5: Search sessions on Mars with 8-minute (left) and 48-minute (right) roundtrip delay: Each point represents a single session with SERP pre-fetching.

search tools from those used by the TREC Session Track, not all pages from the SERPs appear in our top k. In reality, of course, the SERP documents would be the top-ranked subset of our top k, so that exactly k documents would be transferred to Mars.

We analyzed hits on these uploaded documents, where we counted as a hit any topically pre-fetched document that later appears in a SERP from the same session. Having these pages already on Mars potentially allows the Martian user to access them without having to wait for the SERP in which they first appear. With $k = 1000$ we achieve a hit ratio of over 21%; with $k = 2000$ we achieve a hit ratio of over 27%. These hit ratios should translate into substantial reductions in E, although a reasonable estimate requires many assumptions about user behavior, which we avoid in this paper. Unfortunately, this potential improvement to E comes at a big cost to D, as compared to SERP pre-fetching alone, since D is approximately equal to k. That is, we marginally improve effort at a great cost in transferring data that is never used.

Query Suggestions. Most commercial search engines suggest query reformulations and extensions to help guide their users through search sessions. We might take advantage of these suggestions by executing them on behalf of the Martian user, uploading the results and their linked pages, along with the main SERP and its linked pages. If a reformulation by a Martian user matches a query suggestion, we completely avoid a query-response cycle to Earth. Even if the Martian makes an unanticipated reformulation, the additional uploaded information might allow her to continue working while waiting for a response from Earth.

To explore the potential of this idea, we submitted queries from our log to the Bing Autosuggest API,[6] and compared suggestions to queries appearing later in the same session. For 57 queries, a suggested query appeared verbatim later in the same session. While this is less than 2% of all possible queries, it is clear that the idea has some potential, perhaps by going deeper in the suggestions list or by extracting related terms from suggested queries. While some suggestions are spelling corrections or simple morphological variants, others are more complex, e.g., "uglai recipe" for "kenya cooking".

Combining our various query pre-fetching ideas may provide a reasonable overall solution. When a query is received from Mars, we might imagine expanding it with terms from query suggestions, and through other expansion methods, generate a large set of related documents. These could be

[6] `http://www.microsoft.com/cognitive-services/en-us/bing-autosuggest-api`

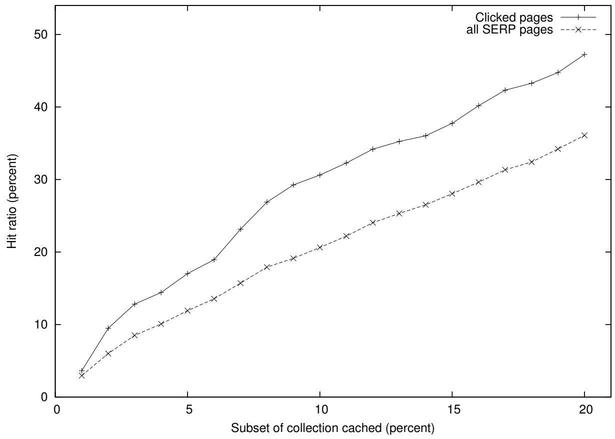

Figure 6: Cache hit ratios for clicked pages and for all SERP result pages.

returned to Mars for re-ranking and exploration by the Martian user. We might even stream documents continuously, similar to Section 4, adjusting the stream on the basis of queries and other feedback from Mars. We leave the investigation of such ideas for future work.

5.3 Caching

As an alternative or in addition to pre-fetching, we could minimize user effort by (partially) caching a snapshot of the web on Mars (we discuss the possible logistics below). If we maintain a partial snapshot on Mars, perhaps we could serve most of the user traffic from that cache, or at the very least give the user some preliminary results to work with while we are fetching the full results. But of course, much of the web consists of lower quality pages that would rarely appear in a SERP, and would even more rarely receive a click. The question, of course, is which parts of the web do we send over to Mars? Caching will greatly increase D, but if the pages are chosen based on some type of static rank, or "page quality", we may be able to reduce E.

As in the previous section, the experiments reported here used the ClueWeb12 crawl and TREC session log. For static ranking, we used the method of Cormack et al. [16], which has performed well on ClueWeb collections (the method is the source of the Waterloo spam scores that are widely used by academic researchers) and has fast code available. Static ranking is based on content only. We trained over the ClueWeb09 collection—an earlier crawl gathered by the same group at CMU in 2009—using as training labels TREC relevance assessments created as part of various experiments using that collection. More specifically, we trained on: 1) all documents judged relevant for any purpose (e.g., for any query) regardless of grade, which were taken as positive examples; 2) all documents assessed as spam, which were taken as negative examples; and 3) a random sample ($N = 3000$) of documents judged as non-relevant, which were also taken as negative examples. The static ranker was then applied to all pages in ClueWeb12. Note that the training data is completely disjoint from this collection, and so there is no "information leakage" from the session data.

Based on this static ranking, we might cache a fraction of available pages on Mars. Figure 6 shows hit ratios for cached pages appearing in the log, considering either all pages linked from SERPs or just pages that were actually clicked. Hit ratios are shown for various caching ratios between 1% and 20% of the full collection. The hit ratio for clicked pages is substantially higher than that for SERP pages generally, helping to confirm the success of our static ranking. By maintaining a 20% snapshot on Mars, we can achieve a hit ratio for clicked pages of nearly 50%.

Location	Lag (min)	Average time (sec)	Average pages transferred	Effort ratio (E)	Data ratio (D)
Earth	0	172.323	3.940	1.000	1.000
Mars	8	445.918	16.351	2.812	5.148
Mars	48	1936.334	16.351	12.561	5.148

Table 3: Average performance for Earth-based and Mars-based sessions under various delay scenarios, with SERP pre-fetching and 20% caching. Average pages transferred and data ratios exclude the 157 million cached pages.

To simulate the impact of caching pages on Mars, we require some assumptions about user behavior in addition to the actual behavior captured in the log. Each session starts as usual, with the user issuing the query appearing in the log. The query is sent to Earth, which follows the SERP pre-fetching approach in Section 5.2, returning the SERP itself and all pages linked from the SERP that are not already on Mars. Meanwhile, the query is also sent to the local cache, which we assume returns a SERP covering the pages in the cache. The user interacts with this local SERP until the log shows she would have clicked on a result not present in the local cache. At that point our simulated user waits for the full Earth-generated SERP before proceeding.

If the user issues further queries, we follow the same process, sending the query to Earth and allowing the user to interact locally without delay. Delay occurs only when the log shows a click on a result not present in the local cache. While the Mars-based user would not actually be able to click on non-local results, since they would not appear on the locally-generated SERP, we take the click as a signal of dissatisfaction with the local results. Since we have no way of knowing how the real user would have proceeded, waiting for the Earth-based results provides a worst-case assumption that places an upper bound on E.

Figure 7 plots the results of this simulation with 20% caching. Here the y axis shows only pages transferred beyond those already cached. Compared to Figures 4 and 5, overall session times are substantially reduced, although they are still well above the Earth-based times in Figure 3. Average performance appears in Table 3; average pages transferred and data ratios exclude the 157 million cached pages.

How might we actually cache a snapshot of the web on Mars? While in our simulations, 20% of the collection represents a "mere" 157 million pages, 20% of *the entire web* remains a substantial, even daunting amount of data. The most practical approach is to physically transport the cached data on cargo rockets (i.e., a sneakernet). The problem, of course, is the transit time: many of the pages will already have changed by the time the cache arrives at Mars. Physical transport of data needs to be accompanied by updates sent from Earth—which of course consumes valuable bandwidth. Without updates, searchers on Mars would be interacting with a copy of the web that is several months old.

The combination of sneakernet and incremental updates frames an optimization problem that commercial web search

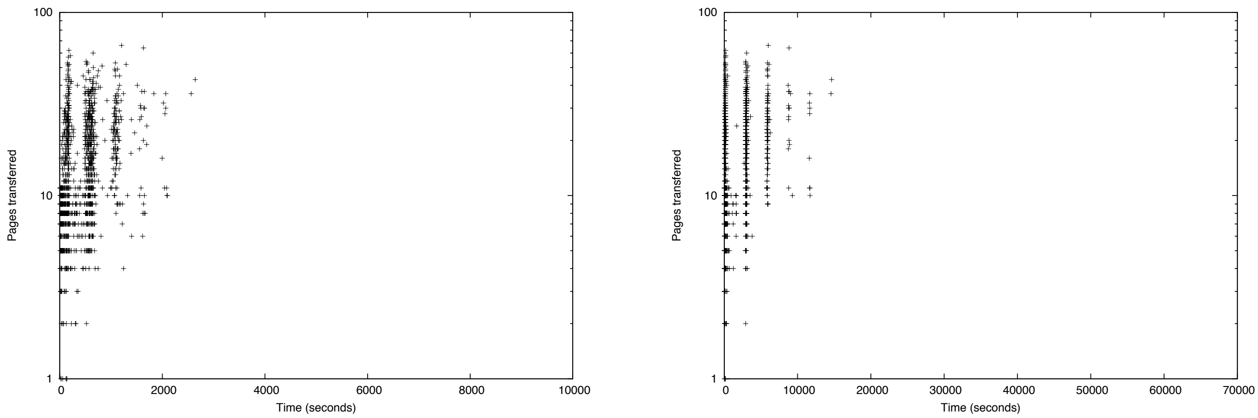

Figure 7: Search sessions on Mars with 8-minute (left) and 48-minute (right) roundtrip delay: Each point represents a single session with SERP pre-fetching and 20% caching. The range of the x axis is the same as in Figures 4 and 5. Pages transferred exclude pages in the cache.

engines are equipped to solve. Today, they must decide what and how frequently to recrawl existing content, and as a result have detailed historic data indicating which pages are "stable" and which pages change rapidly. With this information, it is possible to trade off physical data transport with bandwidth-consuming updates. Although we do not have access to this information, it is a matter of engineering to figure out the best solution. This is a solvable problem.

6. INTERPLANETARY INTERNET

Our two case studies assume the existence of an interplanetary Internet (or at least a Mars-Earth Internet) while ignoring many technical implementation details (e.g., the specifics of delay-tolerant networking). In addressing possible solutions to searching from Mars, the case studies have made differing assumptions regarding caching and interactions. To better harmonize these results, a full exploration of interplanetary WWW infrastructure might proceed in parallel from the bottom up, with the definition of standardized application frameworks and caching policies, and with full consideration given to interactions between potential colonies. For example, a Jovian colony on Europa might preferentially request data from Mars, rather than Earth, when the alignment of the planets place Mars at closer proximity. A reasonable goal of such frameworks would be relative agnosticism from the underlying data transport protocol. The design of this infrastructure represents the primary starting point for our future work.

As long as interplanetary colonization remains in a state of nascency, infrastructural development will require a simulation testbed. To facilitate the development of realistically-deployable solutions, this testbed must account for planetary orbits, rotations, occlusion due to the sun, relay satellites, power disruptions, and all such factors. The testbed must support both real-time simulation, providing a UX designer with the direct experience of using the web from Mars, as well as providing accelerated simulation of various scenarios, such as the impact of a solar coronal mass ejection. Creation of the simulation infrastructure represents a substantial undertaking distinct from the creation of the interplanetary WWW infrastructure itself.

7. CONCLUSIONS

In this paper, we provide a framework for evaluating search from Mars as a tradeoff between "effort" (waiting for responses from Earth) and "data transfer" (pre-fetching or caching data on Mars). The contribution of our work is articulating this design space and presenting two case studies that explore the effectiveness of baseline techniques. Our simulations suggest that we can trade off effort against data transfer to varying degrees. While this paper explores only simple techniques, they set the groundwork for future studies (e.g., caching head queries on Mars).

As we noted earlier, the problem of searching from Mars has analogs closer to home. Instead of a cache we ship to Mars on a cargo rocket, we might FedEx hard drives of web data[7] to rural villages in the developing world, where the village elders can plug these caches into the central wifi access point. This shared access point can intercept web searches with the local cache; usage log data can determine the pages that arrive on next month's hard drive shipment. This scenario parallels exactly search from Mars, and thus searching from Mars is more than idle blue-sky speculation. Furthermore, the breakthroughs that are needed—for example, better session models and models of long-term user needs—stand to benefit web search in general.

Moving forward, we are exploring the problem of supporting access to a broader range of web services (e.g., transactional sites, dynamic content), as well as social media (e.g., push notifications) on Mars. In the medium term, we hope to build the simulation framework discussed in Section 6. Our goal is to create a fully-tested and ready-to-go solution for use by future colonists.

Exploration is perhaps one of the most innate human drives. While we typically associate space exploration with rocket scientists, structural engineers, and geologists, permanent colonies also require services that allow their members to maintain contacts with friends, family, and society. Developing appropriate infrastructure requires careful consideration and experimentation, with this paper representing one small step in that direction.

[7]In fact, this is how the ClueWeb collections are actually distributed.

8. REFERENCES

[1] Delay-tolerant networking architecture. https://tools.ietf.org/html/rfc4838. Accessed: 2017-01-31.

[2] Disruption tolerant networking for space operations. https://www.nasa.gov/mission_pages/station/research/experiments/730.html. Accessed: 2017-01-31.

[3] How does the Mars base communicate with Earth? http://www.mars-one.com/faq/technology/how-does-the-mars-base-communicate-with-earth. Accessed: 2017-01-31.

[4] How long does it take to get to Mars? http://www.universetoday.com/14841/how-long-does-it-take-to-get-to-mars/. Accessed: 2017-01-31.

[5] Time delay between Mars and Earth. http://blogs.esa.int/mex/2012/08/05/time-delay-between-mars-and-earth/. Accessed: 2017-01-31.

[6] B. Aldrin. The call of Mars. *New York Times*, June 2013.

[7] D. M. Boroson. Overview of the Lunar Laser Communication Demonstration. *ICSOS*, 2014.

[8] B. Carterette, A. Bah, E. Kanoulas, M. Hall, and P. Clough. Overview of the TREC 2013 Session Track. *TREC*, 2013.

[9] B. Carterette, E. Kanoulas, M. Hall, and P. Clough. Overview of the TREC 2014 Session Track. *TREC*, 2014.

[10] J. Chen, L. Subramanian, and J. Li. RuralCafe: Web search in the rural developing world. *WWW*, pp. 411–420, 2009.

[11] C. L. A. Clarke, G. V. Cormack, J. Lin, and A. Roegiest. Total Recall: Blue sky on Mars. *ICTIR*, pp. 45–48, 2016.

[12] G. V. Cormack and M. R. Grossman. Evaluation of machine-learning protocols for technology-assisted review in electronic discovery. *SIGIR*, pp. 153–162, 2014.

[13] G. V. Cormack and M. R. Grossman. Autonomy and reliability of continuous active learning for technology-assisted review. *arXiv:1504.06868v1*, 2015.

[14] G. V. Cormack and M. R. Grossman. Multi-faceted recall of continuous active learning for technology-assisted review. *SIGIR*, pp. 763–766, 2015.

[15] G. V. Cormack and M. R. Grossman. Engineering quality and reliability in technology-assisted review. *SIGIR*, pp. 75–84, 2016.

[16] G. V. Cormack, M. D. Smucker, and C. L. A. Clarke. Efficient and effective spam filtering and re-ranking for large web datasets. *Information Retrieval*, 14(5):441–465, 2011.

[17] E. Howell. SpaceX's Elon Musk to reveal Mars colonization ideas this year. *Space.com*, January 2015.

[18] E. Kanoulas, B. Carterette, M. Hall, P. Clough, and M. Sanderson. Overview of the TREC 2012 Session Track. *TREC*, 2012.

[19] E. Kanoulas, P. Clough, B. Carterette, and M. Sanderson. Overview of the TREC 2010 Session Track. *TREC*, 2010.

[20] E. Kanoulas, M. Hall, P. Clough, B. Carterette, and M. Sanderson. Overview of the TREC 2011 Session Track. *TREC*, 2011.

[21] D. D. Lewis, Y. Yang, T. G. Rose, and F. Li. RCV1: A new benchmark collection for text categorization research. *Journal of Machine Learning Research*, 5, 2004.

[22] J. Lin, C. L. A. Clarke, and G. Baruah. Searching from Mars. *Internet Computing*, 20(1):77–82, 2016.

[23] A. S. Pentland, R. Fletcher, and A. Hasson. DakNet: Rethinking connectivity in developing nations. *IEEE Computer*, 37(1):78–83, 2004.

[24] A. Seth, D. Kroeker, M. Zaharia, S. Guo, and S. Keshav. Low-cost communication for rural internet kiosks using mechanical backhaul. *MobiCom*, pp. 334–345, 2006.

[25] M. D. Smucker and C. L. Clarke. Time-based calibration of effectiveness measures. *SIGIR*, pp. 95–104, 2012.

[26] J. Teevan, K. Collins-Thompson, R. W. White, S. T. Dumais, and Y. Kim. Slow search: Information retrieval without time constraints. *HCIR*, 2013.

[27] W. Thies, J. Prevost, T. Mahtab, G. T. Cuevas, S. Shakhshir, A. Artola, B. D. Vo, Y. Litvak, S. Chan, S. Henderson, M. Halsey, L. Levison, and S. Amarasinghe. Searching the World Wide Web in low-connectivity communities. *WWW*, 2002.

[28] G. H. Yang, M. Sloan, and J. Wang. *Dynamic Information Retrieval Modeling*. Synthesis Lectures on Information Concepts, Retrieval, and Services. Morgan & Claypool Publishers, 2016.