

Lessons from El Paquete, Cuba's Offline Internet

Joel Lenin Pinargote Bravo, Rafael Beto Mpfumo, Luis Alejandro Madruga Milanés,
Ximena Michelle Cueva, Gretel García Gómez, Amalia Gómez Marcheco,
Alberto Fernández Oliva

Universidad de La Habana

Havana, Cuba

JLPB_8@hotmail.com, rafaelbmpfumo@gmail.com, Impiri78@gmail.com,
ximena.cueva.u@gmail.com, gothwin@gmail.com, abeagomez@gmail.com,
afdez@matcom.uh.cu

Jeanna Neefe Matthews

Clarkson University/Data and Society

Potsdam, United States

jnm@clarkson.edu

Sam P. Kellogg

New York University

New York, United States

sk6642@nyu.edu

ABSTRACT

Cuba has one of the highest rates of literacy and education in the world, but also one of the lowest rates of direct Internet access. As a result, Cubans have developed a unique and robust offline system for distributing digital content through removable storage media like USB thumb drives and portable hard drives called El Paquete Semanal (The Weekly Package). In this paper, we describe and analyze the contents of El Paquete and how those contents vary over time and between distributors. We compare it to a less-popular state sponsored alternative called La Mochila (The Backpack) and discuss their respective distribution systems. We also compare these offline content distribution systems to available online access points such as those in public WiFi parks, tourist hotels, and universities.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from Permissions@acm.org.

COMPASS '18, June 20–22, 2018, Menlo Park and San Jose, CA, USA
© 2018 Copyright is held by the owner/author(s). Publication rights licensed to ACM.

ACM 978-1-4503-5816-3/18/06...\$15.00
<https://doi.org/10.1145/3209811.3209876>

Finally, we reflect on how this system, developed in the Cuban context, could be useful in other environments.

Author Keywords

Cuba, Internet, Offline Content Distribution, Digital Media, Networking

ACM Classification Keywords

H.5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous

INTRODUCTION

Internet access rates in Cuba are up substantially from 23% in 2011 to roughly 40% in 2016 [13][14]. However, this still remains one of the lowest rates of direct Internet access worldwide, and the lowest rate in the Western Hemisphere. Most Cubans access the Internet through outdoor public WiFi parks where access is expensive (approximately \$1 USD per hour of access from an average Cuban salary of \$30 per month), inconsistent, and inconvenient in comparison with persistent mobile access. Some are also able to access the Internet at work or through educational institutions. A 2015 article estimated that roughly 5% of Cubans have Internet access at home, and though this figure has certainly gone up since then, at-home access remains rare [15].¹

¹ It is worth nothing that Internet-access rates in Cuba are increasing rapidly and 2017 saw the increased availability of nauta-hogar, ETECSA's

In each of these contexts, bandwidth is constrained and variable, and some content may be inaccessible. As a result, Cubans have developed a unique and robust offline system for distributing digital content through removable storage media like USB thumb drives and portable hard drives called *El Paquete Semanal* (The Weekly Package).

RELATED WORK

The majority of the authors of this paper are faculty and students in Cuba at the University of Havana. We have read with interest a number of recent papers on WiFi parks in Havana [1] and the Cuban Street Network (SNET) [2]. In both cases, we were pleased to see interest in the innovative ways Cubans are connecting to each other and the world. We feel we could provide updates and additions to both of these works based on our own experiences in Havana, but in this paper, we are focusing on another crucial part of the Internet connection landscape in Cuba, specifically the distribution of content through USB-connected storage devices like hard drives and thumbdrives. Earlier this year, Dye et al. published a qualitative study of the human infrastructure behind *El Paquete* based on interview and observation data [24].

EL PAQUETE SEMANAL

El Paquete Semanal is a collection of digital material distributed since approximately 2008 in the unofficial Cuban marketplace as an Internet alternative. It has been called Cuba's Google, YouTube, Netflix, Hulu, and Spotify – all without the Internet [7][8], but these comparisons fail to do justice to *El Paquete*'s particularities and strengths. *El Paquete*, updated and distributed every week, contains a diverse collection of digital content, downloaded from abroad and produced within Cuba, movies and TV shows including new releases, music and music videos, video games, sports programming, digital magazines, books, religious materials, mobile and PC applications, and entire copies of websites including Wikipedia and Revolico, a Cuban

Craigslist-style marketplace. *El Paquete* is popular and widely known throughout Cuba. Interestingly, it is not known as *El Paquete* in all parts of Cuba—for example in Matanzas, only 98 km from Havana, it is known instead as *El Cargue* or *The Load*.

As shown in Figure 1, *El Paquete* is realized very simply as a set of top-level directories, each named for the themed content it contains, such as “Deportes” or sports content, “Deportes [HD]” or sports content in high definition, “Programas TV Americana” or American TV programs, “Games,” “Humor,” etc.

The entire package, assembled each week, is roughly 1 TB in size and contains content in 40-60 top-level directories. The exact list of top-level directories varies week to week and between distributors.

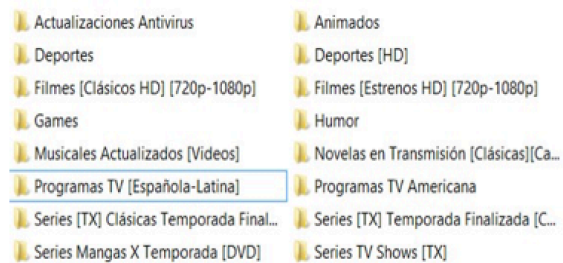


Figure 1. Screenshot of one example of the top-level directory structure of *El Paquete*.

Buyers must provide their own storage medium for the content and hard drives are more expensive than USB thumb drives (*memorias*). It is common to purchase only specific portions of the full content, e.g. choosing specific TV series or movies to fill the purchaser's available capacity. Within Havana, a standard price for a complete copy of *El Paquete* is 2 CUC, or roughly \$2 USD, and less for smaller quantities of data. For specific files, particularly video files (movies, TV, and music videos), it is common for distributors to have multiple resolutions available for purchase, with high-resolution video costing more than low-resolution.

The distribution of *El Paquete*, while not officially sanctioned by the government, has clearly been permitted by way of a lack of government disruption. One reason it is so widely

(the state telecommunications company) home Internet service. We look forward to updated figures.

tolerated is that it does not contain content that is anti-government, subversive, obscene, or pornographic, following the same "no politics, no pornography" policy common within Cuba [2][7]). Much of the content is under copyright in other countries and would be considered pirated material, but this is socially acceptable and legal in Cuba, where intellectual property laws and norms are markedly different.

Almost everyone in Cuba has a source for receiving El Paquete, whether it is a store-front with signage, a less formal neighborhood distributor they pay, or a friend who gives them a copy free of charge [8]. One common manner of purchasing El Paquete is through distributors that physically bring El Paquete to customers' houses on an external hard-drive. Once data is purchased from a distributor, it can then be shared with others e.g. through hand-to-hand transmission on USB storage, over the SNET or over Zapyra or Bluetooth between mobile devices. Thus, while a tremendous amount of content is paid for directly, many Cubans also rely on friends and family members to get content second-hand.

While El Paquete is widely available throughout Cuba, it is not as widely known who is behind its high-level distribution. A string of reporters, artists, academics, and other interested parties have interviewed people at various levels of the distribution chain who have described the processes of production and distribution [5][7], and some people involved in the process are quite open about their roles, goals, business models [7] [12].

Some people specialize in the downloading and/or curation of certain genres of content such as movies, sports, or music. They serve as editors or curators of sections of content, realized simply as a file folder within El Paquete. The means of production or acquisition of content varies widely between editors and across different types of content.

Some curators use higher-bandwidth connections at institutions like universities or tourist hotels to download content from abroad. Others record material from satellite-based connections (which

are often illegal), and still others rely on the passing of physical hard drives from outside of Cuba [8]. The heterogeneity of the contents' origins is sometimes revealed in the content itself: sports programs might have channel watermarks, and movies and TV shows sometimes bear the logo of the downloading team in the corner.

Beyond foreign content, there is also a significant amount of content produced specifically for dissemination on El Paquete. Some examples include Vistar (an independent magazine devoted to Cuban culture), PlayOff (a magazine about sports and sports culture), Hablan2D (a show which previews content from the upcoming week's Paquete), !Seccion Cristiana (Christian religious content), and !!!Sección A R T E (a collection of Cuban art) [8] [9]. El Paquete offers a means of distribution for this custom content that would be unlikely to reach such a wide audience otherwise [8]. In turn, the topology of the network and the means of accessing the content shape the forms that the content crafted particularly for El Paquete takes. For example, a single week's "!!!Sección A R T E" might contain a diverse smattering of e-books and PDFs, audio lectures, video clips, images, and saved HTML documents, taking full advantage of the digital affordances of the distribution system.

Each week, content curators assemble their content sections and pass those sections to a high-level distribution team known as a matriz, which in turn assembles the full package. The finalized package is then passed on to mid-level distributors, who pass it on to other smaller distributors and so on, eventually reaching consumers.

The collection received by an end-user typically passes through several mid-level distributors, particularly for end-users outside of Havana. Starting in Havana, El Paquete content is distributed to all the provinces of Cuba through a system of hard drives passed via car, plane, train, and bus. Mid-level distributors may modify the contents, especially to remove content not of interest to their customers and thus reduce the size of their own distribution, or add content, for

example, local content of interest in parts of Cuba outside of Havana.

Overall, the system of collection, production, and distribution is incredibly timely, with new movies and TV shows appearing in El Paquete and distributed throughout Cuba within days of the content's debut in other countries [16]. This efficient distribution network is rooted in older physical content distribution networks established over decades to disseminate foreign novels and magazines in the 1970s and Beta/VHS/DVD movies in the 1990s and 2000s [9]. This is a clear example of what scholars of historical infrastructural development have termed path dependency, where new networks are layered or built upon older forms (e.g. the tendency of telephone cables to be strung along old railway lines) [23].

Among the most prevalent high-level matrices are Odisea, Omega, and DeltaVision [9][10]. They are often identified by brand-like names and logos that are either included as a video clip or image file within a directory or embedded within the content itself. DeltaVision is the matrix that is most widely known outside of Cuba because of the interviews granted by Danys Cabrera who heads the organization. Within Cuba, however, DeltaVision is a relatively minor player in comparison with Odisea and Omega, much larger organizations in terms of distribution that are nonetheless quieter about their operations [7][9][12].

It is notable that matrices also add advertisements to some content, and it is reported that income from these advertisements rivals income from fees paid by users for the content itself [8]. The general lack of private advertising in Cuba, both on television and in public spaces, makes the development of a small advertising industry surrounding El Paquete content particularly significant. Matrices also play a role in sponsoring and encouraging the production of unique content, exclusive to their version of El Paquete. In this sense, matrices play a variety of roles ascribed to traditional production studios and TV channels, taking care of distribution,

selling advertising and sponsoring the production of exclusive content.

Although monetization is far from universal throughout the network, matrices appear to have a viable and sustainable business model, pending any significant changes in law and governance on the island. However, those involved in content distribution understand the potential fragility of their business model. Changes to government policy, either to crack down on El Paquete or to broaden Internet access, could put the system in jeopardy and/or decrease demand. Similarly, copyrighted material in El Paquete could become a serious point of contention were international intellectual property laws or global business interests to gain more traction in Cuba.

LA MOCHILA

In December 2016, La Mochila (the backpack) was introduced as an officially sanctioned alternative to the extremely popular Paquete. However, unlike the case of El Paquete, La Mochila's means of production and distribution are well known. The Mochila Blog describes the team responsible for producing La Mochila [3]. It is distributed freely through Joven Clubs (youth clubs), a system of over 600 state-sponsored computing clubs throughout Cuba.

While El Paquete has a wide variety of content (Cuban as well as foreign; cultural, religious, and educational), its primary focus is on entertainment, particularly recent foreign content. La Mochila has foreign entertainment, but the primary focus is on cultural and educational content. There are 13 top-level sections with a more polished interface than El Paquete's simple list of directories. Somos el mundo (We are the world), is content from outside Cuba; Me dicen Cuba (Tell me Cuba) is material produced inside Cuba. De mi terruño (From my Native Land), contains productions of the provincial television and community telecentres throughout Cuba, giving increased visibility to productions from outside Havana. Educación para todos (Education for all) is educational material including the Navegante and Multisaber collections and review exercises to help students prepare for the national

college entrance exam (la Enseñanza Superior). Both El Paquete and La Mochila contain applications for mobile devices and PCs, including games. It is worth noting that both El Paquete and La Mochila contain copyrighted content from abroad; the absence of copyrighted material is not the distinguishing factor in the state-sponsored version. La Mochila is simply a smaller state-sponsored alternative to El Paquete with additional national content added.

ANALYSIS OF EL PAQUETE CONTENT

We collected a weekly version of El Paquete from the same distributor in Havana throughout December 2017 and January 2018. Table 1 summarizes the details of this collection. The seven weekly samples vary in size from 855 GB to 931 GB and contain between 9196 and 14447 files. As shown in Figure 1, each weekly sample is organized as a set of top-level directories or categories (43-47 in each of our samples). Some of these top-level directories are repeated each week while others are occasional. Overall, in our 7 samples, there are 89 unique top-level directories. 32 of these occur in all 7 samples; 43 occur only once².

The lower-level structure under each top-level directory varies with some having a deeper underlying directory structure than others. In our analysis, the depth of a file is the number of parent directories between it and the top-level directory. For example, the file “Series TV Shows [Ingles]\ Jimmy Fallon 2017 12 01 Queen Latifah WEB x264 TBS mkv mp4.mp4” is directly under

² The list of the 32 top-level directories which occur in all 7 samples is long, but interesting and illustrative of the content categories even without a translation. They are, in alphabetical order as they appear (note the ! character which pushes the item to the top of the list, and which is particularly characteristic of special folders that do not occur each week): !Sitios de Anuncios Clasificados,!TV Cubana, Actualizacion Antivirus, Animados, Aplicaciones Android, Aplicaciones IOS, Aplicaciones PC, Combos, Concursos de Participacion, Deporte, Documentales, Doramas Finalizados por Temporadas, Doramas en DVD, Doramas en Transmision, Filmes Alta Calidad [AVI,MP4,MKV], Filmes Clasicas [HD], Filmes Estreno [HD],Filmes por Genero Clasicos,Games,Humor [HD], Interesantes, Miniseries, Musica Actualizada MP3 [Exclusiva], Musicales Videos Clip [Exclusiva], Novelas [Clasicas En Transmision], Novelas [En Transmision], Series En Transmision, Series Mangas, Series TV Shows [Español], Series TV Shows [Ingles], Series [Clasicas En Transmision], Shows Latinos.

the top-level directory “Series TV Shows [Ingles]” that contains English TV Series and has a depth of 1. The maximum depth we observed across our samples was 7 and the average depth across our samples was 2.76.

Date	Top-Level Directories	Files	Size (GB)
Dec 4 2017	44	14447	927
Dec 11 2017	47	9196	922
Dec 18 2017	44	10856	931
Jan 1 2018	47	12063	908
Jan 8 2018	46	11512	855
Jan 22 2018	45	10196	924
Jan 29 2018	43	9767	929

Table 1. Summary of El Paquete samples collected by date. Here we report the number of top-level directories, the total number of files and the total size in GB.

We also divided the content we collected into categories by the type of file regardless of the top-level directory in which they are found: video files, audio files, images, documents, Android applications, iOS applications, compressed files, and other. We placed each file into a category based on its file extension. In Table 2, we present a breakdown of content type in the collection, both by size and by number of files.

Video content clearly dominates at 91% of the content by size. However, since video files are large, they represent a much smaller percentage of total files, 31.4%. Images are the largest category by number of files (31.7%), but they represent only 0.3% of the content by size. Excluding compressed archive files (e.g. zip files), which are themselves a collection of many internal files, the next largest category by size is audio files, at 1.6% (space) or 17.3% (number of files). Documents (e.g. PDF files, text files, etc.) are 0.7% of the content by size and 12.2% of the files.

Content Type	Size (GB)	Files
Video	5824 (91.1%)	24540 (31.4%)
Audio	100 (1.6%)	13538 (17.3%)
Images	16 (0.3%)	24757 (31.7%)
Documents	42 (0.7%)	9534 (12.2%)
Android Applications	34 (0.5%)	692 (0.9%)
iOS Applications	11 (0.2%)	128 (0.2%)
Compressed Archive Files	124 (1.9%)	720 (0.9%)
Other	246 (3.8%)	4128 (5.3%)
TOTAL	6396	78037

Table 2. Summary of El Paquete samples collected by content type. Content type is determined by the file extension.

Android applications are more numerous than iOS applications, mirroring the fact that Android-based phones are more common in Cuba than their Apple-made counterparts. PC applications are part of the “Other” category because they represent an even smaller percentage.

We also look at how the file content changes week to week. We took an MD5 checksum of each file and used that to identify files with identical content. Figure 2 reports our results across all 7 samples. As an example, between December 4 and December 11, only 1274 of the 14447 files (8.8%) are identical. These 1274 files represent only 36.23 GB or 3.9% of the 927 GB total content in the December 4 El Paquete sample. If we compare the first December 4 sample to the final January 29 sample, the repeated content has dropped to only 376 files (2.6%) and 1.27 GB (0.1%). Thus, while many of the top-level directories remain the same, their content changes drastically each week: each week’s package is almost entirely composed of brand new content.

REPEATED CONTENT						
	11-Dec	18-Dec	1-Jan	8-Jan	22-Jan	29-Jan
4-Dec	1274 36 Gb	1823 22 Gb	1332 10 Gb	924 1 Gb	666 3 Gb	376 1 Gb
11-Dec		1809 49 Gb	1127 10 Gb	875 1 Gb	520 6 Gb	384 2 Gb
18-Dec			2546 17 Gb	1558 7 Gb	735 6 Gb	505 2 Gb
1-Jan				2421 40 GB	1102 11 Gb	535 5 Gb
8-Jan					1485 14 Gb	909 5 Gb
22-Jan						1241 13 Gb

Figure 2. Repetition of content across our 7 samples. We report both the number of files that are the same and the size of the content that is the same.

We also examined the changes in top-level directory names. In this case, we looked not only at our 7-week sample, but also consulted the year-long archival collection exhibited at Queen’s Museum by artists Julia Weist and Nestor Siré. Their collection included 46 weekly package samples that they collected between August 2016 and August 2017 (roughly 3 months before our samples). The majority of their samples (42 of the 46) were acquired directly from Omega, one the largest high-level distributors, [9] and the remaining four from a different distributor labeled Studio PKT³. We focused on variations in the names of top-level directories in the earliest and latest samples from both Omega and Studio PKT.

Their August 1 2016 Omega sample contained 51 top-level directories while the August 21 2017 Omega sample contained 55. 28 of them are exactly the same. In the August 1 2016 Studio PKT sample, there are 46 top-level directories and in the August 7 2017 Studio PKT sample, there are 50 top-level directories. 28 of them are exactly the same. Between Omega and Studio PKT, however, there only 4 that are exactly the same.

³ Weist and Siré explained that the 4 samples from Studio PKT were acquired from a local distributor to replace missing samples from Omega. Unlike samples directly from Omega, it is difficult to determine the provenance or even the number of levels in the distribution chain for the Studio PKT samples. The same is true for our set of 7 samples. Top-level directory names may be changed and content added/removed at each step in the distribution path.

Of greater interest is the variation between top-level directory names that are not identical. The vast majority of differences are simple predictable variations such as the addition/subtraction of a quality marker like [HD] (e.g. “Humor” vs “Humor [HD]”), the addition/subtraction of one or more leading “!”, the addition of a qualifier for a genre of content (e.g. “Películas por Género [HD] [Fantástico]” or specific actor (e.g. “Películas por Actores {Gwyneth Paltrow}”) or year (e.g. “Trailers [2016]”) or even just a difference between plural and singular (e.g. “Deporte” vs. “Deportes”).

We note other small variations both between distributors and within a distributor. Some distributions appeared to use English terms more widely (e.g. “Games” vs. “Juegos”). Some distributions appeared more careful about correct accent marks in Spanish than others (e.g. “Transmisión” vs. “Transmision” or “Música” vs. “Musica”). However, even within a distributor that tended to be more consistent about accent marks, we observed the same word spelled with and without accents.

Another interesting inconsistency is related to the use of grouping punctuation marks such as {}, (), and []. There did not seem to be a consistent meaning of each type. Even within the same distributor, we saw the use of all three for designating sub-categories of “Películas por Género” (Movies by Genre). For example, from the same distributor, we saw all of the following variations: “Películas por Género (Cine de Época)” with parentheses, “Películas por Género [HD] {Comedia}” with braces, and “Películas por Género [Aventura]” with brackets.

While it is difficult to know if these inconsistencies are accidental or deliberate, their presence suggest that humans rather than algorithms remain central to the process of compiling and naming El Paquete each week. The varied roles people play in the distribution of El Paquete is explored through participant interviews in [24]. It is worth noting that such variations would make it more difficult for others to copy or combine collections programmatically.

Distributors might want to consider the benefits of programmatic control that standardization of names and directory structure would enable.

COMPARISON TO ALTERNATIVES

To put our discussion of El Paquete into context, we summarize the alternatives for both offline and online transmission/consumption of digital content in Cuba in Table 3. Offline access options include El Paquete and La Mochila, as we have discussed in this paper, as well as the Cuban Street Network (SNET) which, although not connected to the global Internet does provide networked communication within communities in Cuba [2] including access to downloaded/cached copies of Internet content. The online access options include WiFi parks [1], access from ETECSA offices and cyber-cafes, tourist hotels, Internet access points at job and education sites, Internet at home (ETECSA’s nauta-hogar service), and cell-phone based (mobile) Internet.

It is worth noting that most Internet access in Cuba (including that available in WiFi parks, hotels, and in homes), is paid for per unit of *time* rather than per bit, leading to different consumption practices and different network affordances. Cell-phone-based Internet like Digicel is sold per bit, but is primarily intended for tourist use, is very expensive, and is not easily available for purchase from within Cuba.⁴

It is also worth noting that acquiring Internet access, either through the purchase of single-use access cards or through a grant of accessibility at a place of employment, is linked to identification documents. For example, purchasing a single-use Internet card at an ETECSA branch usually requires an official form of identification, although it is also a common practice for both tourists and locals to purchase access cards on the street illegally near public WiFi hotspots in urban areas.

We originally added a bandwidth column to Table 3 and populated it with results of testing the

⁴ According to its website, the Digicel will be discontinued on April 9th, 2018. It remains to be seen what kind of services will be offered in replacement.

speed of upload and download. However, our measurements varied widely in all cases. Access speeds at the University of Havana varied widely, but were generally higher than at either the hotels or WiFi parks. We hope to undertake a longer-term study, but from our experience, we expect to see the largest impacts on available bandwidth from changes in time of day and weather given that WiFi parks are accessed outside and vulnerable to the elements.

It is not only access speeds that vary widely. It is also the ability to connect to the network at all. It is not uncommon visit a WiFi park to access the Internet and find yourself either completely unable to connect or able to connect sporadically, rendering websites and services unusable. This can make it difficult for Cubans to coordinate a time to connect with friends and family abroad reliably, and makes light-weight chat applications such as IMO, which use less bandwidth than others, especially popular. This inconsistency also makes it difficult to determine if an inaccessible site is inaccessible due to transient technical difficulties having to do with traffic or infrastructure failure or because of deliberate content blocking on the sending or receiving side (the Apple App Store is blocked by Apple on the U.S. side, for example, so that the company does not break embargo laws). There is a general belief that few sites are actively blocked by the government, but given the cost and technical difficulties, most people treat the Internet access like a scarce resource and focus their online time on only the most important tasks. Accessing potentially-blocked sites to check if they are accessible is typically low on the priority list.

While it is not the same type of speed measurement, it is worth noting that the time to copy El Paquete/La Mochilla data over USB can be substantial. USB 2.0 can theoretically transfer data at 480 Mbps and USB 3.0 over 10 times faster at 5gbps. However, the majority of USB flash drives or USB attached thumbdrives don't deliver anywhere close to these theoretical bounds and write speeds are lower than read speeds [17]. Copying 1 TB via USB can easily

take hours. To put this in perspective, while getting a copy of La Mochilla is free at a local Joven Club, it does require that there be a computer available for making the copy and also requires either staying during the entire copying process (2-4 hours) or leaving an expensive hard drive with the staff who are busy with other tasks and may be unable to monitor the copying process.

LESSONS FOR OTHER ENVIRONMENTS AND REGIONS

As we have described, El Paquete represents one very successful means of addressing a particular set of constraints and challenges that confront people in Cuba today. We are aware of data distribution systems on storage media that exist in other countries as well. Notable examples include [21][22] in North Korea and [25][26][27][28] in India. Garg et al. propose Postmanet, a general system that achieves high bandwidth transfers with a combination of postal mail and storage media [29]. We are most struck with the similarities in [28] including the degree to which users are strongly motivated by the desire to be entertained.

For traditional network-based communications to be effective, there must be a reliable path from the source, through intermediate routers to the user. In regions where electricity and Internet infrastructure is inconsistent or vulnerable to failure—disaster-prone areas and active conflict zones offer examples—the reliability of these paths cannot always be counted upon. Islands, in particular, can be especially vulnerable to disconnection during disasters. The earthquake that hit Haiti in 2010 and Hurricane Irma's destructive effects on Puerto Rico in 2016 offer two recent examples of situations where a formalized hand-to-hand distribution system built upon established social networks and supported by off-grid solar or battery power could potentially have better weathered the storm. Social ties have always proved more flexible and resilient than immobile power and telecom grids.

Offline access models like El Paquete also offer an alternative to the online tracking/surveillance models that are common in other regions. Offline

access provides for one type of anonymity of access – more visibility to people close to you in the distribution system, but less surveillance from government or corporate entities.

One of the authors had an opportunity to propose and discuss the possibility of USB-based distribution of CS research and education materials within the Nigerian academic community. We observe that many offices and academic labs in Nigeria have banks of batteries to weather inevitable drops in available electricity. Similarly, a collaborative system of downloading and sharing of materials could allow students and faculty to weather the variability of Internet access and bandwidth. We specifically proposed that attendees at a yearly computer science conference actively bring USB-drives with downloaded material to share. This would add the sharing of USB-stored content to an already established network.

We also note that a wide variety of educational materials such as MOOC platforms are difficult to operate in completely offline mode. As future work, both for users in Cuba and for other users around the work, we would like to investigate modifications that would allow these materials to work properly offline. While closing the digital divide is usually thought about strictly in terms of offering Internet access, Cuba's case proves that building an expensive fiber-optic backbone is not the only means of attempting to do so. One great challenge to this model is that many established Internet services, particularly social networks, make assumptions about the hardware and infrastructure they run on and, as in the case of many MOOC platforms, will simply not function without a persistent connection. One step that application builders can take now is to ensure that services function offline, with an option to upload, download, or transfer content or communications once a connection is made with the Internet or another local network like SNET.

FUTURE WORK

There are so many ways we would like to extend this work. First, we will make available a longer

analysis of this data with substantially more graphs than we have room to make available in this conference version, including a more quantitative description of La Mochila. Second, we would like to substantially extend our data collection to include for example, copies of El Paquete from other provinces in Cuba and especially look for instances where specialized content is added during distribution. Third, we would like to do a more extensive comparison of data in El Paquete to data available from other sources. For example, we could compare the contents of Wikipedia available within El Paquete to the version online to see if anything is added or removed. Similarly, Revolico is available online but blocked inside Cuba. We would like to identify differences in the El Paquete version and the online version. Fourth, we would like to do some manual classification of content. We are especially interested in marking each file as foreign content or Cuban content and in marking content that is copyrighted. More difficult but potentially interesting would be marking content in categories such as religious or educational. A more systematic study of when content is first made available online vs. when it first becomes available in El Paquete would also be interesting.

CONCLUSION

We have presented a study of El Paquete, Cuba's popular USB-based system for data distribution. We described its complex and efficient distribution network (built on existing networks) which offers a model that other regions could emulate. We compare El Paquete to the state-sponsored alternative, La Mochila. We collected and analyzed 7 weeks of El Paquete samples. Some key findings are that video dominates the file content, the vast majority of content changes each week. Small changes in the set of top-level directories indicate an opportunity for increased programmatic control. El Paquete could serve as a model for other regions in which electricity and Internet access is limited or variable.

	Description	Cost	Features/Limitations
OFFLINE ACCESS OPTIONS			
El Paquete	Offline Internet. Not officially permitted but tolerated and widely available. Need computer to view stored content.	Reusable Storage Device. 2 CUC or roughly \$2 for a full week's package.	Effective replacement for several categories of Internet content especially high BW entertainment focused content. Some content not permitted (e.g. no political content or pornography). Anonymous consumption and private consumption possible.
La Mochila	Offline Internet. Officially supported.	Reusable Storage Device, Free content. Can view content using computers at Joven Club.	Not an Internet replacement; Small collection of materials. Anonymous consumption and private consumption possible.
Cuban Street Network (SNET)	Local network disconnected from wider Internet but with rich content available. Need computer to join network.	Donation to help maintain infrastructure (e.g. 1 CUC or donation of equipment)	Clear codes of conduct (e.g. no political content or pornography). Not anonymous consumption. Conduct monitored.
ONLINE ACCESS OPTIONS			
WiFi Park	WiFi access in outdoor public parks. Need computer or mobile phone to connect. Purchase cards (nauta cards) at ETECSA office or reseller.	1 CUC or roughly \$1 USD per hour	Often connection not successful or BW insufficient. Difficult to use in inclement weather or at night. Use is in public. Some content blocked or at least hard to tell if content is blocked or just access problems. Not anonymous consumption. Specific card number linked to identity documents on purchase.
ETECSA office	Some ETECSA offices or navigation centers allow people to surf the Internet using computers they provide.	Don't need to own computer or removable media. 1 CUC or roughly \$1 USD per hour.	Similar to WiFi Park, but use is indoors during limited business hours. Cheaper access available for a collection of Cuban sites. Monitors in the office.
Hotel-based WiFi	WiFi access in tourist hotels. Purchase cards (nauta cards) at hotel. Not interchangeable between hotels or in WiFi parks.	4-5 CUC or roughly \$4-5 USD per hour	Use in lobby is still relatively public, but more comfortable and protected from weather. 4-5X more expensive than WiFi park. Unknown if some content is blocked. Not anonymous consumption. Specific card number linked to identity documents on purchase.
University-Based/Job-Based Internet	Internet access is available to professionals as a part of their job (e.g. at universities and government facilities)	Free to students/faculty/workers. Computer labs or bring your own device.	Not anonymous consumption. Web browsing through a proxy tied to user id. Some content restrictions (e.g. social media blocked during working hours).
Internet at Home	Nauta Hogar is being set up throughout the country. It is accessible to all and ETECSA makes a telephone call to each family when the service is available in their area. Need computer to connect.	The price is a little cheaper than in WiFi parks up to 30 hours per month and then you can pay at the same price as in parks.	Dedicated connection. 1 MB/s advertised speed. Not anonymous consumption.
Cell-phone based Internet access	Digicel and others offer roaming cell data that can be used within Cuba to access the Internet wherever cell service is available.	100 MB/ \$25 300 MB/ \$50	Cost extremely high. SIMs not generally available for purchase inside Cuba. Unknown if some content is blocked.

Table 3. Comparison of access methods for Internet-based content in Cuba.

REFERENCES

1. Michaelanne Dye, David Nemer, Laura R. Pina, Nithya Sambasivan, Amy S. Bruckman and Neha Kumar. 2017. Locating the Internet in the Parks of Havana. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '17)*, 3867-3878. <http://dx.doi.org/10.1145/3025453.3025728>
2. Eduardo E. P. Pujol, Will Scott, Eric Wustrow and J. Alex Halderman. 2017. Initial Measurements of the Cuban Street Network. In *Proceedings of the 2017 Internet Measurement Conference (IMC '17)*, 318-324. <https://doi.org/10.1145/3131365.3131395>
3. Mochila Blog. <http://mochilablog.cubava.cu>.
4. Revolico: Anuncios clasificados en Cuba. <https://www.revolico.com/>.
5. "El paquete semanal Comes to New York". 2017. <http://www.cubanartnews.org/news/el-paquete-semanal-comes-to-new-york/6457>
6. "La Mochila", alternativa cubana a películas y series piratas. 2017. <https://elcomercio.pe/tecnologia/actualidad/mochila-alternativa-cubana-peliculas-series-piratas-402158>.
7. This is Cuba's Netflix, Hulu, and Spotify – all without the internet. 2015. <https://www.youtube.com/watch?v=fTTno8D-b2E>.
8. Sarah Kessler. In Cuba, An Underground Network Armed With USB Drives Does The Work Of Google And YouTube. 2015. <https://www.fastcompany.com/3048163/in-cuba-an-underground-network-armed-with-usb-drives-does-the-work-of-google-and-youtube>.
9. Julia Weist and Nestor Siré. 17.(SEPT)[By WeistSiréPC]™. September 7 2017 to February 18 2018. Queen's Museum.
10. April Clare Welsh, Give Me Future: Exposing the DIY pirate network feeding Cuba with music . February 27 2017. <http://www.factmag.com/2017/02/27/cubas-diy-pirate-network-el-paquete-semanal/>.
11. Mario Luis Reyes. Detrás del Paquete semanal, Danys. November 17 2016. <https://oncubamagazine.com/ecos/detras-de-el-paquete-semanal-danys/>.
12. DeltaVision. 2018. <http://conocecuba.com/delta-vision/>.
13. Internet World Stats, <https://www.internetworldstats.com/stats2.htm>.
14. Statista, Percentage of population using the internet in Cuba from 2011 to 2016, <https://www.statista.com/statistics/739036/internet-penetration-cuba/>.
15. Emilio San Pedro. 2016. Cuba internet access still severely restricted. BBC News. <http://www.bbc.com/news/world-latin-america-35865283>.
16. Anna Cristina Pertierra. 2012. If They Show Prison Break in the United States on a Wednesday, by Thursday It Is Here: Mobile Media Networks in Twenty-First- Century Cuba. *Television & New Media*, 13(5) 399–414, SAGE. DOI: 10.1177/1527476412443564.
17. Lincoln Spector. 2014. USB 3.0 speed: real and imagined. *PC World*, June 26 2014. <https://www.pcworld.com/article/2360306/usb-3-0-speed-real-and-imagined.html>.
18. Internet Access in Cuba – Where to Find Wifi [2018]. *Internet* <https://asocialnomad.com/cuba/internet-access-in-cuba/>.
19. Rich Potter. The Communications Blockade. *Submarine Telecoms Forum* 58. July 2011. http://www.subtelforum.com/issues/STF_58.pdf
20. Xinhua. Cubans to gain access to mobile internet in 2018. *XinhuaNet*. http://www.xinhuanet.com/english/2017-12/31/c_136862657.htm
21. Flash Drives for Freedom. <https://flashdrivesforfreedom.org/>

22. Joshua Rhet Miller. 2017. Importing hope into North Korea, one USB drive at a time, New York Post, March 17 2017.
23. Marie-Laure Djelic and Sigrid Quack. 2007. Overcoming path dependency: path generation in open systems. *Theor Soc* (2007) 36:161–186. DOI 10.1007/s11186-007-9026-0.
24. Michaelanne Dye, David Nemer, Josiah Mangiameli, Amy S. Bruckman and Neha Kumar. 2018. El Paquete Semanal: The Week’s Internet in Havana. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI '18). ACM, New York, NY, USA, Paper 639, 12 pages. DOI: <https://doi.org/10.1145/3173574.3174213>
25. Neha Kumar and Nimmi Rangaswamy. 2013. The mobile media actor-network in urban India. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI '13). ACM, New York, NY, USA, 1989-1998. DOI: <https://doi.org/10.1145/2470654.2466263>
26. Jacki O'Neill, Kentaro Toyama, Jay Chen, Berthel Tate, and Aysha Siddique. 2016. The Increasing Sophistication of Mobile Media Sharing in Lower-Middle-Class Bangalore. In *Proceedings of the Eighth International Conference on Information and Communication Technologies and Development*(ICTD '16). ACM, New York, NY, USA, Article 17, 11 pages. DOI: <https://doi.org/10.1145/2909609.2909656>
27. Nithya Sambasivan, Ed Cutrell, and Kentaro Toyama. 2010. *ViralVCD*: tracing information-diffusion paths with low cost media in developing communities. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI '10). ACM, New York, NY, USA, 2607-2610. DOI: <https://doi.org/10.1145/1753326.1753721>
28. Thomas N. Smyth, Satish Kumar, Indrani Medhi, and Kentaro Toyama. 2010. Where there's a will there's a way: mobile media sharing in urban india. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI '10). ACM, New York, NY, USA, 753-762. DOI: <https://doi.org/10.1145/1753326.1753436>
29. Nitin Garg, Sumeet Sobti, Junwen Lai, Fengzhou Zheng, Kai Li, Randolph Y. Wang, and Arvind Krishnamurthy. 2005. Bridging the digital divide: storage media + postal network = generic high-bandwidth communication. *Trans. Storage* 1, 2 (May 2005), 246-275. DOI=<http://dx.doi.org/10.1145/1063786.1063791>