Taking Benefit from the User Density in Large Cities for Delivering SMS

Yannick Leo, Anthony Busson, Carlos Sarraute, Eric Fleury

Abstract

In this paper, from a measurement study and analysis of SMS based on traces coming from a nationwide cellular telecommunication operator during a two month period, we propose two DTN (Delay Tolerant Network) like network protocols for delivering SMS. More precisely, we perform a temporal and spatial analysis of the Mexico City cellular network considering geolocalized SMS. The temporal analysis allows us to detect events and to check for overloading periods, with abnormal or unexpected traffic, and to study the evolution of classical parameters such as activity or distance between source and destination. The spatial analysis is based on the Voronoï diagram of the base stations covering Mexico City. We explain how SMS traffic can be characterized. Such key characterization allows us to answer the question: is it possible to transmit SMS using phones as relay in a large city such as Mexico City? We defined two simple network protocols to transmit SMS from a source to a destination. These DTN like protocols do not need routing nor global knowledge. These protocols take benefit from the locality of SMS, the density of phones in Mexico City and the mobility of phone users. We studied a mobile dataset including 8 millions users living in Mexico city. This gave use a precise estimation of the average transmission time and the global performance of our approach. After 30 minutes, half of the SMS were delivered successfully to destination. These protocols are shown more efficient during the rush hours as nodes density and mobility are higher. Moreover, it is the precise moment that the infrastructure network needs to be relief.

1 Introduction

The need of communicating in a dense city is always increasing. Every day, millions of SMS are sent in a large city like Mexico City. Traditional SMS is challenged by alternative messaging services such as Facebook Messenger, WhatsApp [1], Tango, Skype and Viber by using data connexion and/or wireless hot spot [2]. One can also use P2P applications to enable the connection of smartphonesvia Bluetooth or Wi-Fi without an Internet connection. In this case, people have to constitute a dense network to insure the connectivity. Though it was not designed in purpose, FireChat was used as a communication tool in some civil protests. Nevertheless, SMS is still a growing market and remains a

very popular service over cellular networks since 82% of mobile users are sending SMS in our dataset. The SMS is well-known and well-used in both developed and developing countries.

During rush hours, SMS traffic may consume a non negligible part of the backbone network capacity, and sometimes saturates it. This saturation may come from the SMS architecture itself which is totally centralized. Every SMS is deliver to a unique SMS CENTER (SMSC) which acts as a centralized, store-and-forward server that is responsible for accepting, storing, retrieving subscriber information, and forwarding SMS to the intended destination of the SMS. It is becoming a great challenge to increase the amount of traffic delivered to the users while keeping the infrastructure stable (i.e., same number of relays and backbone capacity). Apart from the typical rush hours, the mobile network can be globally challenged during special events such as natural disasters or locally saturated during concerts, conferences, riots or sport matches. Mobile networks are dimensioned to sustain the load 99% of the time, but for those specific events, the activity can be incredibly higher than the congestion threshold. In [3], authors revealed that voice calls and SMS are still in use in these large scale events, despite frequent reports from users about the data network unavailability. It can thus be interesting to propose and test new protocols, less dependent of the cellular infrastructure and/or the backbone, that could carry a part of the traffic load.

In this study, we evaluate the benefits and the feasibility of a delay tolerant network (DTN) approach to transmit SMS and more generally data from a source to a destination. Instead of using classical routing, we use relays close to the source and phone users that are connected to those local relays to reach the destination. The advantage of our approach is that we do not perform a routing algorithm, we do not need global knowledge, neither its associated mechanisms such as neighbor discovery, exchange of control messages, etc. and do not need to know where the destination is. Moreover, as we only use local relays that are close to the source, the bandwidth cost of a SMS is smaller and the backbone infrastructure of the operator is not used. On top of that, our protocol works best when the capacity of the network is challenged during rush hours as the density of phones and the mobility of users are even higher.

It is important to notice that we clearly do not target an implementation of our approach in existing 3GPP standards or existing cellular networks protocols. One goal is to demonstrate that the DTN approach could be feasible and helpful. We demonstrate by replaying millions of real SMS traffic the reliability and the gain of our proposals. Note that such approach could be used in future cellular standards, and also be used as a key enabler for P2P applications.

The contributions of this paper are two-fold. We perform an analysis in time and space of SMS traffic. It is based on a significant database describing SMS sent during two months in Mexico city [4]. It represents 90 millions of phone users all over Mexico. The temporal analysis includes an event detection to check overloading periods and the evolution of the SMS activity at different time scales. The spatial analysis is based on the distance between source and destination in kilometers and number of hops. The number of hops is computed

according to the Voronoï diagram of the base stations. Thanks to the traffic characterization and more precisely its regularity and locality, the second contribution is the proposal of two protocols that aim to carry SMS traffic, and to relieve the infrastructure network in terms of load. We use the same real SMS traces to evaluate the efficiency of these less centralized protocols. Through the replay of these SMS, we show that with a very simple and not optimized algorithm, it is possible to have a delivery ratio higher than 50%. It is obviously not perfect, but we consider that it is enough to prove the benefit of our approach.

The paper is organized as follows. We first reference studies that are linked to our experiments and topic in Section 2. In Sections 3 and 4, we present our large trace that contains SMS and localized calls and we perform a spatial and temporal analysis to point out some pertinent characteristics of the cellular networks. The two protocols are presented in Section 5. In Section 6, we show the performance evaluation of these two protocols in terms of transmission success and delay, and point out how much higher activity and mobility can increase the sucess rate of our protocols. We conclude in Section 7.

2 Related work

A study [5] of nationwide cellular network analyzes the temporal activity and peaks. It seems to be a great advantage to understand better the activity variations to limit the congestions. The overload during rush hours or specific events that induces high traffic is a great problem and diverse solutions have been proposed [6, 7]. It is possible to limit the size of the message, to include in mobile phone a software that reacts to congestions, or even to try to recognize and predict congestions. Other studies expose an application built on top of SMS which enables to use SMS for several services. In [3], the author proposes a reliable SMS-based data transport protocol. These applications use the ubiquity, reliability and the possibility of building SMS-based applications to adapt the classic cellular protocol.

Yet some events are unpredictable, such as natural disasters or riots. Both global and local events can be unpredictable. Many studies propose general or specific methods to detect such events [8, 9, 10]. During these overloaded periods while the network is saturated, some messages are not delivered unless they use alternative ways of communicating [1], or more decentralized applications such as the BitTorrent's P2P Encrypted Messaging App where messages are directly sent from one phone to another without being routed through or stored on any servers.

Sms communication is now used for many services and the list of uses continue to increase. It is now possible to use SMS to make health campaigns [11, 12, 13], schedule calls, conduct electronic surveys, provide e-voting services, send calendar notifications, search the Internet, and exchange status updates with servers on the Internet.

We now understand better how people are communicating with their mobile phone in terms of activity [5] and mobility [14, 15]. In this paper we analyze

3

two aspects which are linked to the locality of SMS and the density of users. The observation of certain periods of great density and mobility of phone users combined with the locality of SMS and a high activity lead us to create two protocols that take advantage of these characteristics.
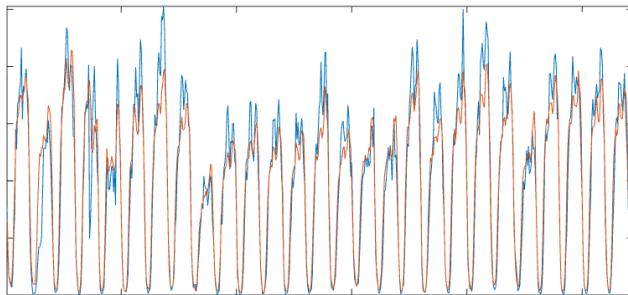
3   Data source

Figure 1: Global SMS activity per hour in Mexico City for 28 days over 2 months (in blue) and the geolocalized sampling for which the locations of the source and the destination are known ofr the same 28-day period (in red). These curves are normalized by the mean.

We used traces extracted from mobile phones [4] representing 92 millions of clients in Mexico and 8 millions in Mexico City during a 2-month period: from March to April 2014. The city is covered by 775 base stations that are part of the telecommunication network of a cellular operator. This anonymized dataset contains 70 millions of SMS and 170 millions of phone calls all over Mexico. Some calls are localized, *i.e.,* we know the base station of the source or destination. From these mobile calls, it has been possible to localize in Mexico City 1.5 millions of SMS for our study. For each SMS, we set the localization of the closest call in time. We assume that if the source and the destination received or made a localized call 30 minutes before or after the SMS, we can effectively set a source and destination location for the SMS. We also noticed that more than 92% of geolocalized SMS sent from Mexico City have the destination in Mexico City.

For our study, we define a graph with the base stations as nodes, for which there is a link/edge between two nodes if they are neighbors in the Voronoï tessellation: there is a link between two base stations (bs_1, bs_2) if and only if bs_1 and bs_2 have a common border in the Voronoï tessellation built according to the base station locations. Consequently, we can define a distance between two base station not only in term of kilometers but also in terms of hops through this graph. Two neighbors are at distance 1, the neighbors of the neighbors are at a distance 2, and so on. As we do not have any hierarchical information on base stations, we consider that every base station has the same role.

4

Validation As we take a sampling of SMS for which we have the locations of the source and the destination, it is important to test that our sample is representative of the global trac. In Figure 1 we show the global activity and the sampled activity. Both are normalized by their mean. It seems at a first glance that both signals of the activity per hour are similar. The Pearson correlation coefficient of the two time series is equal to 0.95, a linear equation describes the relationship between both signals perfectly. As the p-value associated is null, the computed correlation is significant. We also comfirm this validation by a cross-correlation of two discrete-time signals, it shows a strong similarity between both signals.
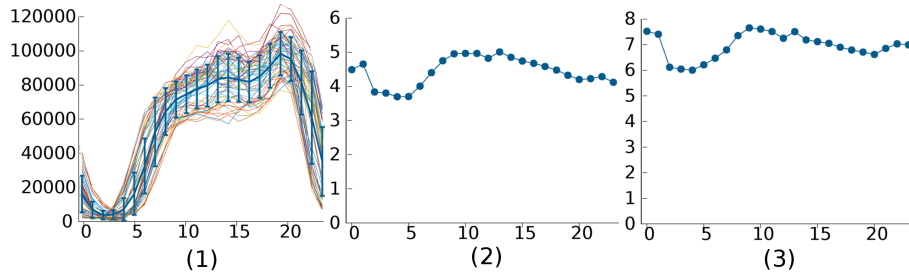
4 Data analysis

Figure 2: For each hour of the day, we show the average of three parameters over the 2-month period: (1) average and standard deviation of the number of SMS (bold blue) and distributions for all days (muticolor); (2) average distance in hops of SMS; (3) average distance in km of SMS.

In this section, we make an analysis of the localized SMS according to three parameters: the activity (number of SMS sent), the distance in kilometers and the distance in hops (distance in hops between source and destination of the SMS in the cellular graph, defined by Voronoï diagram, from source base station and destination base station). These parameters are computed according to the time (per hour) and the space (per cell). The temporal analysis consists in studying the number of SMS and the distance of SMS per hour. To deeper analyze the traffic patterns, we use an empirical mode decomposition (EMD) method. Using the EMD method, the time series can be decomposed into a finite and small number of components. These components form a complete and nearly orthogonal basis for the original signal. Without leaving the time domain, EMD is adaptive and efficient. Since the decomposition is based on the local characteristic time scale of the data, it can be applied to nonlinear and non-stationary processes. Once the time series is decomposed, we are able to detect traffic behavior that does not fit the general periodicity and extract the high frequency component that gather abnormal events. The spatial analysis

5

is based on the Voronoï diagram from base stations, and aims to show up the spatial diversity.

Temporal analysis and event detection As we expected, the activity varies through time at several scales. During the day (from midday to 8pm), the activity is greater than during the night. When we average the activity on a daily base for each hour or when we superimposed activity for every days (Figure 2), we observe a period of lower activity during the night and higher activity during the day. The standard deviation is quite small and variations are periodical. Although the density of mobile phones varies during the day, it also varies between different days. For instance, the activity during weekdays is greater than during the weekend, and the typical peak is reached on Friday at 6 pm.

Moreover, the activity sometimes increases because of specific events. In order to detect and validate such observations and have a better insight on the time series representing the SMS activity per hour, we perform an Empirical Mode Decomposition (EMD) on the signal [16]. The signal is the number of SMS per hour during two months. The EMD allows to represent the non-stationary signal as sum of zero-means Intrinsic Mode Function (IMF) and one residue. Figure ?? gives the decomposition of the global SMS activity. The IMF 3 clearly gives a daily periodic signal (a spectral analysis also give an harmonic decomposition in days of the signal). Let observe the high frequency IMF 1 plotted on Figure ??. We plot in red the mean of the IMF 1. The signal is clearly oscillating around the mean in a compact envelope with few extra peaks of activity. In green, we plot the mean plus twice the standard deviation and extract all points greater than this value in order to point out specific events. For example, in our trace, this signal shows some perturbations, the biggest one happened on the 18th of April 2014 (day 47 on the figure) just after an earthquake near the Pacific coast of Mexico occurred in the state of Guerrero, 265 km southwest of Mexico City.

We also studied the distance in km and in hops that quantify the proximity between source and destination. As one can observe in Figure 2, these distances are not varying a lot through hours but during the night SMS are more local than during the day. One hypothesis might be that people may send SMS at home. To estimate how local the SMS are, we can compare the average distances of our trace to a random case. For instance, the average in km between two random base stations is equal to 16.4 km whereas, in our trace, the average in km of SMS sent is equal to 6.9 km. Similarly, the average of distances in hops in the graph induces by Voronoï diagram is equal to 14.2 for random base stations and 4.7 for localized SMS. The SMS seems to be much more local than the random case. The locality property can be used to deliver SMS without any routing as many destinations are close to the source.

Analysis per base station From an operator point of view, it is interesting to study the traffic in each cell as managed by the base stations. In Figure 3,
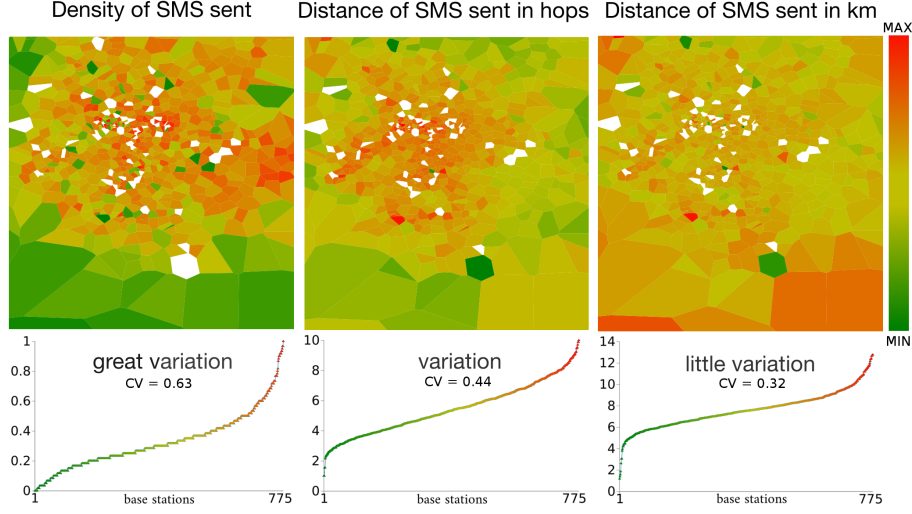
6

Figure 3: For each base station, we show the density of SMS sent which corresponds to the normalized ratio between the number of SMS sent and the size of the cell, the distance in hops of SMS sent and the distance in km of SMS sent. The result is plotted in two different ways: (Top) a heat map to give a visualization of the quantities and (Bottom) quantities according to base stations are computed from the lowest to the highest value, the min/max are given in these figures below. (left) The density of a cell is computed as the ratio of the number of SMS sent for a base station and the area size; (middle) the average distance in hops of SMS sent; (right) the average distance in km of the SMS sent.

we can notice that the distribution of density per base station is heterogeneous as the variation coefficient $CV = \frac{\sigma}{\mu} = 0.63$. Some base stations have a high activity according to the surface whereas others have quite low density. The average of the number of SMS sent is 1378 SMS during that period. It also shows that the density is higher in the city center rather than in suburbs. Although the distance in hops is more homogeneous than the density ($CV = 0.44$), one can notice a larger distance in the city center area. It is certainly due to the spatial density of base stations, which is higher in the city center to manage the higher traffic load. The number of hops to send a SMS at a given distance is then greater. Even if the average of the distance in km per SMS sent for each cell is quite constant ($CV = 0.32$), this averaged distance seems to be a bit longer for the SMS sent from the city center.

5   Protocols

We describe two protocols to deliver the SMS in another way than the protocol used in the cellular network. Our approach relies on the density and mobility
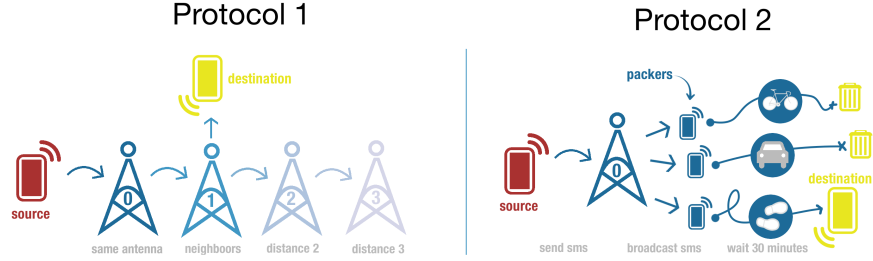
Figure 4: Sketches of the protocols 1 and 2.

of phone users combined with the locality of SMS. The protocols use the base stations that are close to the base station attached to the source and users that are connected to it. They are not totally decentralized but more centralized than the classic cellular protocol that consist in a routing through the base stations. Figure 4 depicts these protocols.

Protocol 1 The source sends the SMS in the usual way to the base station to which it is attached (Figure 4). Then, this base station retransmits to the stations around itself which are neighbors (all the stations that are at a distance 1). This process can be repeated in order to reach the neighbors of the neighbors at distance 2 from the original station. Considering a fixed distance k, the SMS is well received if and only if the destination is attached to a base station at a distance $\leq k$. The efficiency of Protocol 1 depends on the locality of the exchanges. During this process, the location of the destination does not have to be known.

Protocol 2 The second protocol relies on the mobiles attached to the original base station of the sender (Figure 4). When this base station receives a SMS, it duplicates this message to the mobiles which are attached to that station. In practice, every mobile in that cell gets the SMS, they are called the packers. If the destination is one of these packers, the SMS is already transmitted. The packers who have just received the SMS move through Mexico City and may switch from that antenna to others. As the packers are moving, if a packer and the destination are at the same time attached to the same antenna, the destination will receive the SMS. At this moment, the packer sends the SMS to the base station and the base station sends the SMS to the destination as it is attached to this antenna. The success of this communication depends on the density of packers and thus of mobile users and on their mobility. If many packers are moving randomly all over the city, the probability of reaching the destination has to be very high.

6   Results
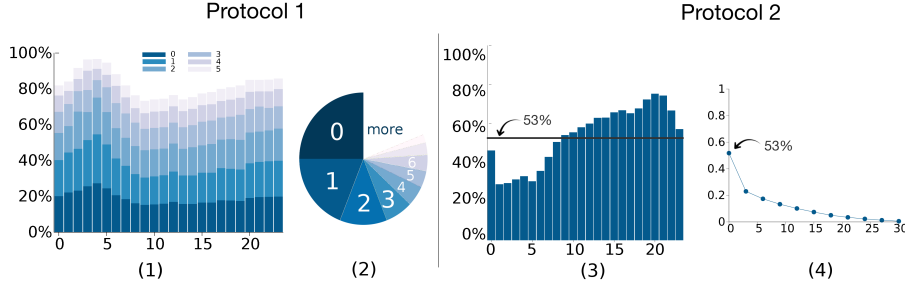

Protocol 1

Protocol 2

(1)   (2)   (3)   (4)

Figure 5: (1) Temporal transmission success for Protocol 1 as function of the hours of the day. We consider different number of hops for the transmission of SMS; (2) global transmission success according to the number of hops averaged over the day. (3) Temporal transmission success for Protocol 2 averaged by day hours where packers keep the message 30 minutes then remove it from their phones; (4) inverse cumulative distribution of the global transmission success according to the delay (≤ 30 min).

We validate our concept by evaluating the feasibility of our approach. It is evaluated through the percentage of SMS that are properly/successfully delivered. We are not competing with classical cellular routing but we show that our protocols can be an alternative for some applications to complement it by relieving the infrastructure network of a part of its load during some challenging time, for instance. For Protocol 2, we choose to limit the delay to 30 minutes which seems to be a reasonable upper limit for delivering a SMS.

Results for Protocol 1   We quantified the efficiency of the first protocol in Figure 5 for which the success is only based on the locality of SMS. We show that 20% of SMS sent reached their destination if we only consider people attached to the same base station ($k = 0$), almost 35% when the base station transmits the SMS to its neighbors ($k = 1$) and 46% if we consider base stations at a two-hop distance ($k = 2$). This result highlights the fact that the destination is usually close to the base station of the source. In comparison, if we consider the base station network of Mexico City, the average of hops between two random stations is 11 and the diameter (maximum distance in hops) of this network is 22. Considering $k \leq 5$, only 13% of the base stations are reachable. In that sense, we can say that there is a high locality of SMS.

Results for Protocol 2   In the traces that we studied for the network of Mexico City (1.5 million SMS over 2 months), one third of SMS are delivered in less than 10 minutes and half of them are transmitted in 30 minutes (Figure 5), this delay is set arbitrarily and reasonably. We consider that after 30 minutes without being delivered, the transmission fails. One can notice that 20% of
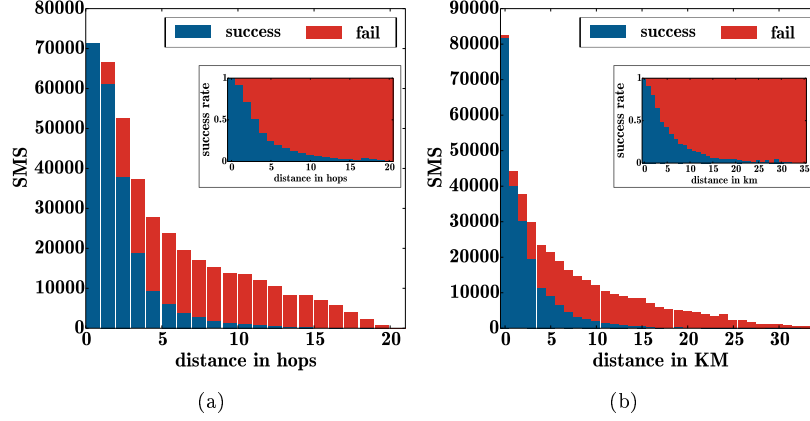
9

(a)　　　　　　　　　　　　　　(b)

Figure 6: Analysis of the success rate of the protocol 2 according to the original distance between the source and the destination. (left) Number of SMS that reaches the destination, number of SMS that are dropped and total number of SMS according to the distance in hops of the SMS; (left, focus) Success rate according to the distance in hops of SMS; (right) Number of SMS that reaches the destination, number of SMS that are dropped and total number of SMS according to the original distance in km of the SMS; (right, focus) Success rate according to the original distance in km. There is the original distance between the source and the destination when the SMS is sent and the delivering distance that corresponds to the distance between the source and the destination when the message is delivered.

transmission success is due to a packer that switch from an antenna to another. The mobility is a key point for that protocol. When the activity is high, during rush hours, the protocol has more than 70% of delivering success. It is interesting to notice that the efficiency of our protocol is the best when the operator service is challenged. Moreover, as we miss out some locations (only 9% of the SMS were localized), we just have a part of the packers and have an underestimation of the efficiency. The protocol 1 relies on the locality of SMS sent during the day whereas the protocol 2 also takes advantage of the mobility and density of phone users leading to a transmission success rate that increases drastically during the rush hours with a maximum rate of 70%.

Intuively, the distance in hops or in km between the source and the destination has a great impact on the success rate. If the source is far from the destination, the packers should be very quick and hopefully reach the destination within 30 minutes. As we can notice on the figure 6 (focus), the success rate is decreasing drastically as the original distance between the source and the destination is increasing. When the number of hops is equal to 0, the success
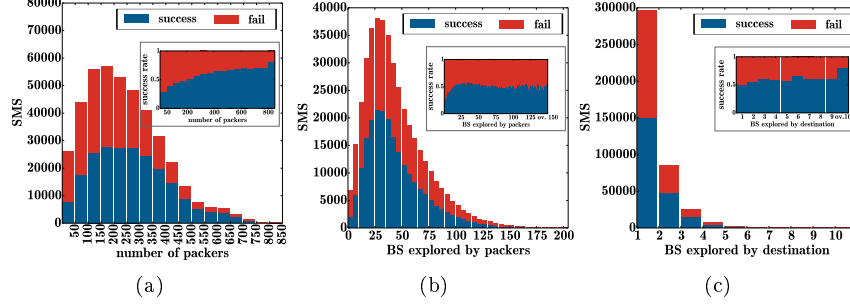
Figure 7: (a) The average success rate, percentage of SMS delivered, according to the number of packers. (b) The average success rate, percentage of SMS delivered, according to the number of base stations reached by the packers during 30 minutes. (c) Average number of explored base stations according to the hours of the day. It represents the average mobility for each hour of the day.

rate is 100% whereas it is null for a maximum distance in hops of 21. The distance in km gives the same results. Yet, SMS are very local and the global success rate is boosted by this local property. Few SMS have a long distance, the social interactions between people in Mexico City are local.

In the protocol 2, the packers have a key role as they deliver SMS by moving along the city. In our dataset, we only have 6% of them. It seems difficult to estimate the underestimation of our success rate due to the lack of geographical information but it seems to be significant. In Figure 7a, we evaluate the success rate according to the number of packers. When the number of packers is very low, the success rate is approximately 40% and increase as the number of packers is increasing. We have to notice that the limit should not be 100% because long distance SMS have a very low probability to be delivered even with more packers. Figure 7a not only shows that our protocol success rate is actually an underestimation but it gives us an undirect measure of the mobility of phone users. Like the density of users and the locality of interactions, the mobility is a capital asset for the mobile users ad hoc network. More precisely, we quantify the mobility of packers by counting the number of base stations explored during the 30-minute experiment. For each SMS, we count all the base stations explored by at least one packer. In Figure 7b, we measure the delivering success rate according to the number of base stations explored by the packers. One can intuitively make a link between the mobility of the packers and the amount of base stations explored. When the mobility of packers is very low, the success rate is not suitable, close to 20%. As, the mobility of packers is increasing, the rate success is becoming greater than 50%. The difference between having a very small number of explored base stations (less than 20) and having a decent number of explored base stations (greater than 20) has a great impact

on the delivery success. We also point out in figure 7c that the mobility of the destination has a small effect on the result. When the destination is exploring more than 10 base stations, the rate success is even greater than 80%. Yet, if the destination is not moving, the result are still maintained.

It seems that the number of packers representing the density and their mobility are positively correlated. These two parameters have a direct impact on the success of the protocols. We empirically proove, that the correlation between the mobility and the activity is temporally strong (Figure 8). Users are tending to move and use their mobile phone at the same hours of the day. It shows that we can obtain the full potential of the user adhoc network during rush hours when users are active and are moving.

It is worth noting that we only considered a subset of data traces. The real number of packers is thus higher and the delivering probability is sub-estimated. Nevertheless, even in this case, we have shown that our protocols offer an efficient alternative to the default SMS delivering.
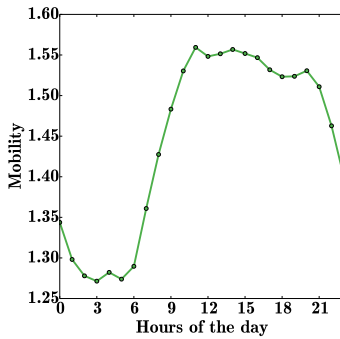

Figure 8: Average number of explored base stations according to the hours of the day. It represents the average mobility for each hour of the day.

7   Conclusions

In this paper, we have proposed two simple protocols based on DTN principles to carry a part of the SMS traffic in a cellular network. They have been evaluated through an original method as we had the opportunity to replay a large trace of geolocalized SMS in Mexico City. This evaluation has shown that the density of the network, users mobility and locality of SMS increases the efficiency of the protocols and so could unload efficiently the backbone network of the operator. Even if there are already some solutions that are close to our protocols, to our knowledge the evaluation of the efficiency of such approaches through the use of huge datasets has never been performed.

The protocol 1 offers a transmission success up to 50% whereas it reaches 70% for protocol 2 during rush hours, which shows the ability of our approaches

to deliver a non negligible part of the SMS and to relieve the infrastructure network of a part of its load. Moreover, we did not used any synthetic model to compute SMS and localizations, our experiments are based on a real traces. We quantified how much the activity and the mobility can improve the success rate of the Protocol 2. In a large city such as Mexico City, the density and the mobility of people is very high. Our experiment reveals that the mobile users ad hoc network has a interesting potential.

Many variations and combinations of these protocols can be proposed and evaluated with this trace. For instance, we decided for the second protocol to set $k = 0$ by sending only to the source base station but it would have been possible to increase k. At the second step, the packers are defined as the clients attached to the source base station. Another possibility is to set a probability of spreading the SMS and to create a more epidemic protocol. When a packer switches from a base station to another, he can transmit the SMS to the new base station which can retransmit to all the clients associated to it. The number of packers will greatly increase. Yet, there is a trade-off to find between our protocols and the broadcast protocol as the SMS has to be stored and transmitted to all the packers and every transmission has a cost. At this point we need to make a compromise between the computational and storage gain we expect and the efficiency of the protocol.

References

[1] K. Church, R. de Oliveira, What's up with whatsapp?: Comparing mobile instant messaging behaviors with traditional sms, in: Proceedings of the 15th International Conference on Human-computer Interaction with Mobile Devices and Services, MobileHCI '13, ACM, New York, NY, USA, 2013, pp. 352–361. doi:10.1145/2493190.2493225.
URL http://doi.acm.org/10.1145/2493190.2493225

[2] C. K. Toh, Ad hoc mobile wireless networks: protocols and systems, Pearson Education, 2001.

[3] H. N. Curto, J. Caetano, J. Almeida, A. Ziviani, C. H. S. Malab, H. T. Marques-Neto, Using sms to transfer small data packets during periods of high workload on mobile data networks, in: XXXIII Simpósio Brasileiro de Redes de Computadores e Sistemas Distribuídos, 2015.

[4] C. Sarraute, P. Blanc, J. Burroni, A study of age and gender seen through mobile phone usage patterns in mexico, in: Advances in Social Networks Analysis and Mining (ASONAM), 2014 IEEE/ACM International Conference on, 2014, pp. 836–843. doi:10.1109/ASONAM.2014.6921683.

[5] P. Zerfos, X. Meng, S. H. Wong, V. Samanta, S. Lu, A study of the short message service of a nationwide cellular network, in: Proceedings of the 6th ACM SIGCOMM conference on Internet measurement, ACM, 2006, pp. 263–268.

[6] E. Oliver, Characterizing the transport behaviour of the short message service, in: Proceedings of the 8th International Conference on Mobile Systems, Applications, and Services, MobiSys '10, ACM, New York, NY, USA, 2010, pp. 223–238. doi:10.1145/1814433.1814457.
URL http://doi.acm.org/10.1145/1814433.1814457

[7] G. Heine, M. Horrer, GSM networks: protocols, terminology, and implementation, Artech House, Inc., 1999.

[8] W. C. Young, J. E. Blumenstock, E. B. Fox, T. H. McCormick, Detecting and classifying anomalous behavior in spatiotemporal network data, in: Proceedings of the 2014 KDD Workshop on Learning about Emergencies from Social Information (KDD-LESI 2014), 2014, pp. 29–33.

[9] V. A. Traag, A. Browet, F. Calabrese, F. Morlot, Social event detection in massive mobile phone data using probabilistic location inference, in: Privacy, Security, Risk and Trust (PASSAT) and 2011 IEEE Third Inernational Conference on Social Computing (SocialCom), 2011 IEEE Third International Conference on, IEEE, 2011, pp. 625–628.

[10] M. Faulkner, M. Olson, R. Chandy, J. Krause, K. M. Chandy, A. Krause, The next big one: Detecting earthquakes and other rare events from community-based sensors, in: Information Processing in Sensor Networks (IPSN), 2011 10th International Conference on, IEEE, 2011, pp. 13–24.

[11] B. S. Fjeldsoe, A. L. Marshall, Y. D. Miller, Behavior change interventions delivered by mobile telephone short-message service, American journal of preventive medicine 36 (2) (2009) 165–173.

[12] T. L. Lai, Service quality and perceived value's impact on satisfaction, intention and usage of short message service (sms), Information Systems Frontiers 6 (4) (2004) 353–368.

[13] C. Déglise, L. S. Suggs, P. Odermatt, Short message service (sms) applications for disease prevention in developing countries, Journal of medical Internet research 14 (1) (2012) e3.

[14] B. C. Csáji, A. Browet, V. A. Traag, J.-C. Delvenne, E. Huens, P. Van Dooren, Z. Smoreda, V. D. Blondel, Exploring the mobility of mobile phone users, Physica A: Statistical Mechanics and its Applications 392 (6) (2013) 1459–1473.

[15] Y.-A. de Montjoye, C. A. Hidalgo, M. Verleysen, V. D. Blondel, Unique in the crowd: The privacy bounds of human mobility, Nature srep. 3.

[16] G. Rilling, P. Flandrin, P. Gonçalves, On empirical mode decomposition and its algorithms, in: Proceedings of IEEE-EURASIP Workshop on Nonlinear Signal and Image Processing NSIP-03, Grado (Italy), 2003.
URL http://perso.ens-lyon.fr/paulo.goncalves/pub/NSIP03_GRPFPG.pdf