Fine-Grained Intercontact Characterization

Tiphaine Phe-Neau
UPMC Sorbonne Universités
tiphaine.phe-neau@lip6.fr

Marcelo Dias de Amorim
CNRS and UPMC Sorbonne Universités
marcelo.amorim@lip6.fr

Vania Conan
Thales Communications
vania.conan@fr.thalesgroup.com

***Abstract*—So far, efforts attempting to characterize the spatiotemporal nature of disruption tolerant networks (DTN) have relied on the dual notion of *contacts* and *intercontacts*. A contact happens when two nodes are within communication range of each other. An intercontact is simply defined as the dual of a contact, i.e., when two nodes are not in communication range of each other. We believe that such a characterization, which we refer to as "binary", is not enough. In real situations, familiar nodes tend to form groups while others seek to avoid each other or display precise meeting patterns. In this paper, we focus on the plethora of situations beyond the binary hypothesis – in other word, we investigate the structural properties of the topology when nodes are not in contact but do have a contemporaneous path connecting them. To the best of our knowledge, this is the first time a paper investigates this issue. We first introduce the notion of *n-ary intercontact* and, to defend its adoption, we quantify the proportion of nodes bearing this new intercontact notion in well-known datasets available to the community. Surprisingly, we observe that most pairs of nodes are nearby (within a few hops) for significant amounts of time when not directly in contact. Finally, we compare the impact of our definition over classic intercontact for DTN characterization and give incentives on using the n-ary intercontact definition to leverage resulting new communication opportunities.**

I. Introduction

With the growing penetration of intelligent devices like smartphones into people's everyday life, disruption-tolerant networking (DTN) emerged as a groundbreaking communication paradigm in the modern networking landscape [1]. To satisfy the communication needs among those devices in an opportunistic way, many new stimulating approaches have been proposed in the literature [2], [3], [4], [5]. A common substrate to these approaches is to rely on the complementary concepts of *contacts* and *intercontacts*. The notion of contact has a factual definition, it is when two nodes are within direct communication range of each other. An intercontact is defined as the complementary of a contact, i.e., simply when two nodes are *not* in contact. As we will extensively investigate in this paper, such a simplistic definition for intercontacts (henceforth mentioned as *binary intercontacts*) ends up being a melting pot for any attempt to properly benefit from the geographic proximity of users.

Fig. I represents a network snapshot illustrating our concerns. From A's point of view, it has two nodes in contact (nodes B and C). With the binary intercontact definition, all four remaining nodes are considered in "intercontact" mode. In such a situation, most DTN approaches infer the impossibility of exchanging messages via multihop paths and often calls for a "wait" period until it meets the destination or find someone

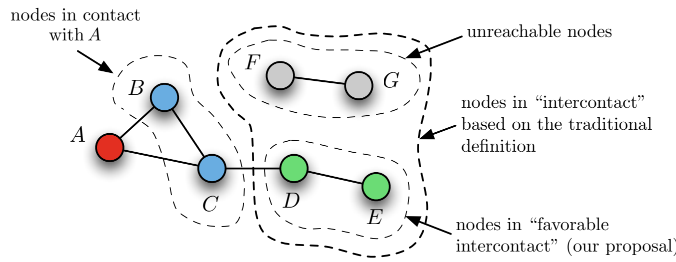


Fig. 1. Node A neighborhood "contact and intercontacts": A is in *contact* with his two direct neighbors, all four remaining nodes are in *binary intercontact*. However, we can reach the lower pair of nodes in intercontact via a contemporaneous path, they are in a *favorable intercontact* state.

else that knows the destination better (based on some other criterion). By denying the inherent ad hoc network part in DTN, we cannot pull the best of both worlds. In our example, we observe that ad hoc communication would succeed if A had data to transfer to either D or E. Contemporaneous paths may exist between nodes not in contact. We refer to this property as *favorable intercontact*. A natural question that arises is: should node A use multihop communications to send data to D and E or wait until it gets within direct communication range with them? This question raises two issues: (i) how to know when two nodes are in the vicinity of each other (when not in direct contact) and (ii) how design of efficient protocols and algorithms to properly exploit such situations.

In this paper, we address the first issue. More specifically, we tackle the fundamental problem of knowing whether a pair of nodes shows sufficient sociability to spend significant share of the time in the vicinity of each other. To this end, we investigate several real-world datasets available to the research community and provide insights into fine details of intercontact periods. We formalize our proposal by defining the notion of *n-ary intercontacts*, where n stands for the distance separating two nodes.

Our work has its foundations on several previous contributions of the literature. As outlined many times, user mobility patterns are not random [6]. People have a tendency to form communities or oblivious groups and then, they display "favorable" intercontact properties. Gaito et al. based their study on workplaces as they also felt the force of social patterns in DTN [7]. There are induced relationships between people-carried devices. Other works, such as the ones of Whitbeck et al. [5] and Borrel et al. [8], propose to classify opportunistic networking into several categories depending on the expected group formation when nodes move around. Our work is complementary to the aforementioned ones as it makes a step further by analyzing *node proximity in a pairwise way* and by identifying *hidden intercontact possibilities* that are often underestimated by traditional forwarding solutions.

We make several interesting observations in our study. Firstly, we confirm that the binary contact-intercontact idea is too rough to efficiently capture potential communication opportunities while nodes are not in contact. Secondly, for a significant amount of pairs, we observe that nodes spend as much time at one and two-hop distances. Thirdly, for some datasets, about half of the traditional intercontact time is in fact a connected. We do believe that our results will motivate protocol designers to first check close neighborhood before adopting a wait strategy.

In summary, the contributions of this paper are:

- New ternary vision for accurate DTN understanding through n-ary intercontact characterization.
- Empirical analysis to show that the binary contact assumption is not enough via an evaluation of binary vs. n-ary intercontact behaviors in existing datasets.
- Extended comparison between binary and pathless intercontact distributions.

II. N-ARY INTERCONTACT: DEFINITION AND OBJECTIVES

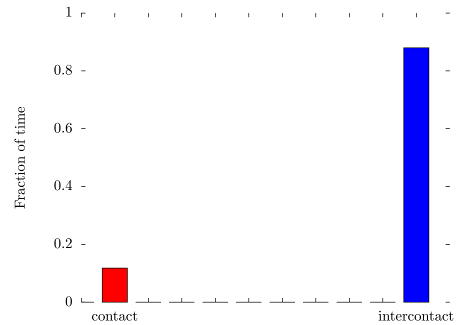
In this section, we provide the necessary background as well as an intuitive explanation of our DTN characterization. We also point out situations where our approach obviously becomes interesting.

A. Pairwise visualization

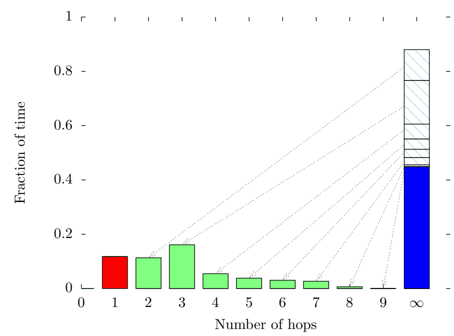
Considering a high level abstraction for intercontact erases punctual events happening between given nodes (for instance, people eating at the same place, at the same time, every week day). Yet, we could use them for efficient forwarding. To maintain traces of these events, we choose to perform pairwise analysis. Bearing in mind the inaccuracy of the binary intercontact vision, we thought of n-ary intercontact as a reflector of human sociostructure as well as ad hoc communication opportunities. We divided it in two intuitive notions: Favorable and Pathless intercontact.

B. N-ary intercontact: an intuitive definition

1) Favorable intercontact: Intercontact is favorable when we have end-to-end transmission possibilities. The n parameter ($\in [|2; +\infty|]$) displays distance between nodes and is a key ingredient in our study. This simple benchmark provides the required information to decide which strategy a protocol should use. Depending on the time spent at a certain distance, we can also derive the surrounding stability. Favorable intercontact can also be seen as extended ad hoc communication chances.



(a) binary intercontact: nodes are in contact for 10% of the time else they are in traditional intercontact.



(b) n-ary vision: like in a) nodes have 10% contact but in reality, they dwell at a distance 2 for around 10%, at a distance 3 for 16%. Real intercontact deprived of multihop path represent only 50% of the timeline (∞).

Fig. 2. Example of Time-Distance distribution (Random Trip dataset)

2) Pathless intercontact: In opposition to favorable situations, Pathless intercontact indicates the lack of end-to-end path between a pair of nodes. This granularity reveals real DTN frameworks and can influence a node's resolution to use asynchronous transmission scheme from delaying transmission to specific routing approaches.

C. Motivating example

In Fig. 2, we generated synthetic Random Trip traces and compared conventional intercontact with our n-ary vision. When a pair of nodes spend 10% of their time in contact (1 hop), with the binary intercontact vision (Fig. 2(a)), we would say that they spend the remaining 90% of their time in intercontact. We have shown in Fig. I that this definition fits for nodes bearing no path between them. But is it enough for nodes displaying multihop connectivity? The traditional intercontact supposition ignores all these possibilities while we could use them for alternative low delay reliable data forwarding.

In Fig. 2(b), with our improved intercontact, we observe that, even though in intercontact, nodes stay at a 2 hop distance during 10% of the experiment duration, at a 3 hop distance for 16% and at a 4 hop distance for 5%. These distances indicate contemporaneous paths through nodes extended neighborhood, hence, new delay less transfer opportunities.

The *Contact, Favorable and Pathless intercontact* triplet is the new ternary classification we suggest for a simple yet meaningful DTN characterization.

D. Potential usages

MANET, DTN or hybrid network. With the impressive literature on ad hoc networks, rejecting the MANET-DTN correlation would be a waste. With a more accurate intercontact understanding, a node may decide to use MANET transmission instead of DTN algorithms or even discriminate between wireless interfaces like Wi-Fi, Bluetooth or 3G. Any protocol has to choose the relevant strategy depending on its current setting. We can even think of new hybrid approaches using ad hoc and disruption tolerant approaches together (see Section V).

Adaptive DTN. Our fine grained intercontact helps discriminating user behaviors. For example, let us consider three nodes A, B and C. We consider node A wants to communicate with another node. When in contact, they use common approaches. But, in intercontact, what can we do? If A is in contact with B on a frequent basis or else completely out of reach, adopting a WAIT strategy (delay transmission until the next encounter) sounds optimal and avoids bandwidth waste. If A never meets C, an usual Spray and Wait [4] strategy looks optimal. From then on, we could consider adaptive DTN strategies with destination-based parameters. They bring meaningful resource savings and optimize every nodes peculiarity. Yet in a melt situation, where A seldom sees D and dwells at a close n distance, we can switch between different types of approaches. We can perform MANET transmission when the situation is stable and D close, an adaptive Spray and Wait for a more fickle setting (spray less copies or wait a little more). There are many parameters an intercontact understanding can change for DTN comprehension.

III. Datasets

To test the n-ary intercontact concept, we choose several datasets often used in DTN characterization studies. Our aim is to study representative behaviors to observe the part of Favorable and Pathless intercontact in every nodes timeline.

Infocom05 is a dataset obtained in a conference scenario [9]. It involved 41 iMotes relating participants motions for around 5 days. We focus on the second day (12 hours) because it has the highest contact observations from the whole dataset. *Infocom05* depicts a working background with sessions, social event and their induced micro commuting. Each iMote performs a scan every 120 seconds.

Infocom06 is another conference-based dataset with 78 iMotes for 3 days [9]. We focus on the second day as it is the only one with full unbiased data. The first day, researchers handed iMotes forcing a gathering and the third day they collected them back throughout the day resulting in partial results for some devices. Other parameters are the same as in *Infocom05*.

Rollernet involved 62 iMotes during a 3 hour Rollerblade tour in Paris [10]. Leguay et al. set a shorter scanning granularity of 15 seconds. We sense in this description a highly dynamic situation with inherent strong social relationships.

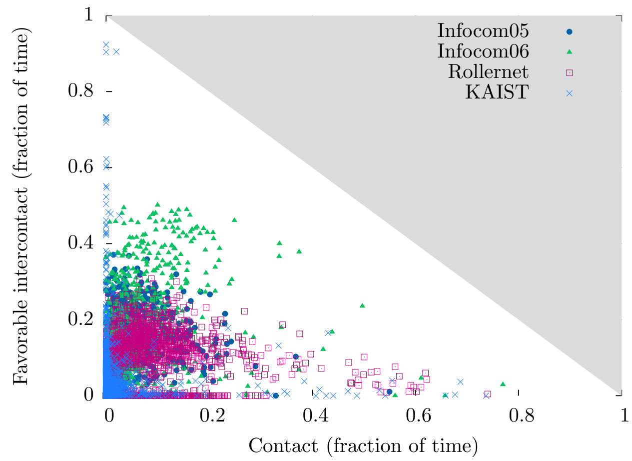


Fig. 3. Contact - Favorable intercontact mapping: for each pair of nodes, we plot the fraction of time they display Favorable intercontact depending on the fraction of time they spend in contact. Most pairs have non null Favorable intercontact opportunities indicating a pervasive property. We also observe a wide variety of pairwise behaviors. This encourages us to define pairwise forwarding techniques instead of an overall tactic.

KAIST and NewYork come from NCSU's study []. Unlike *Infocom05* or *Rollernet*, they are not contact traces but GPS coordinates (Global Positioning System). We converted them via movement simulation and affected every nodes a 10 meters wireless range transmission. We assume each file comes from a specific user. *KAIST* therefore has 92 files for campus-based movements and *NewYork* has 39 in a city wide measurement.

IV. Results

In this section, we present an experimental evaluation of our n-ary intercontact definition. We start analyzing the proportion of nodes bearing Favorable intercontact.

A. Binary or n-ary intercontact?

For each timeline, we extract the fraction of time a pair of nodes stays in contact and the cumulated time they spend in Favorable intercontact. We present these behaviors in Fig. 3. Each dot embodies a pair of nodes behavior. Crosses on the top left of our plot depicts nodes spending very few instants in contact (around 1% of their timeline duration) but bearing Favorable intercontact properties for 90% of their timeline. This indicates these nodes are close with end-to-end paths to one another but hardly ever stay in direct contact. Pure Pathless intercontact is located around the x-axis, these are pair of nodes with real DTN meeting patterns. The triangle in the bottom right of Fig. 3 is an example where nodes remain around 80% of their timeline in contact else they dwell in a Pathless intercontact state.

We first observe that many couples display a non null fraction of time with Favorable intercontact. The percentage of nodes having at least 1% of Favorable intercontact during their timeline duration is 76% for *Infocom05*, around 57% for

Rollernet and *Infocom06* and 29% for *KAIST*. 39% of *Rollernet* nodes and 62% of *Infocom05* have favorable intercontact properties for more than 10% of their respective experiment duration. These transmission possibilities are clearly pervasive. However, with the binary intercontact assumption, we miss all these chances while we could use them to guide routing decisions or actually send data.

The importance of Favorable intercontact properties in these datasets quantifies how inaccurate binary intercontact is for DTN characterization. Its use ends up in massive transmission opportunities losses. In the next section, we notice how our ternary vision gathers oblivious sociostructure information.

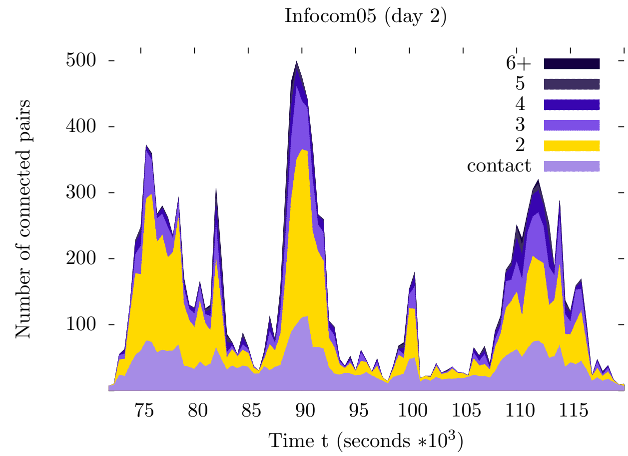
B. Sociostructure and Favorable intercontact opportunities

N-ary intercontact keeps track of the shortest distance between nodes. It embeds precious information about nodes close neighborhood. A given pair of nodes can be either in contact, in Favorable intercontact or in a Pathless intercontact state. Here, we focus on contact and Favorable intercontact information. These two states embed a distance notion. Nodes are in contact if they are close enough to be within each other's range. They are in Favorable intercontact if the density between them is high enough to provide a contemporaneous path to one another. The closer nodes are, the more interests they share, the more likely they are to meet again and the more useful they can be concerning data forwarding.

Couple of nodes can be in contact or linked via one or more relay nodes on their shortest path. From Fig. 4 to Fig. 7, we present what we call an aggregated network sociostructure. There, we plotted the number of connected pairs for each shortest distance. For instance, nodes in layer 2, are connected via a 2 hop paths, they have a node between them acting as a relay. The bottom layer symbolizes the amount of pair of nodes in contact. Above, each layer indicates the sum of nodes connected by a n hop distance or more. We quantify these additional end-to-end transmission opportunities in corresponding tables underneath each plot. The first line indicates the number of neighbors on average for each distance. Below, we display the maximum neighbors for each distance. These transmission opportunities are powerful as they only involve few relays and enable low delay communications.

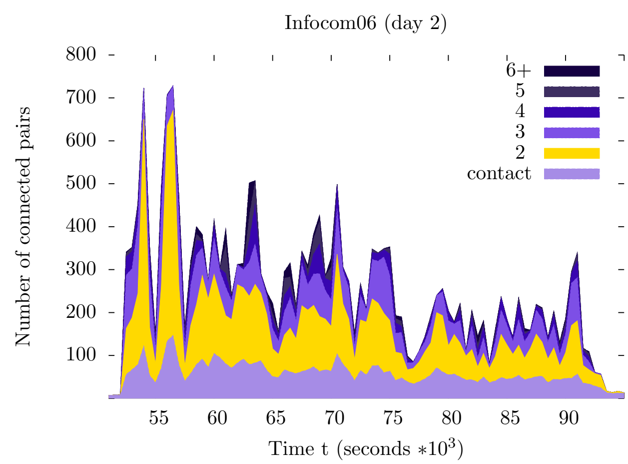
In Fig. 4, we observe several peaks of connected pairs with their extended contact neighborhood. Knowing *Infocom05* is a conference-based measurement, we can correlate them with morning arrivals, lunch, afternoon break and end of sessions. An unexpected visual observation is how distance 2 Favorable intercontacts overcome contact opportunities during these peaks of high density. These environments of highly connected crowds ignite Favorable intercontacts. As a result, Favorable intercontact-based transmissions should be more helpful than direct contact transmissions or real DTN schemes.

For *Infocom06*, another convention dataset, time division are less clear. However, we maintain our observation concerning 2 hop distance connected pairs (see Fig. 5). As a rule, distance 2+ Favorable intercontact overthrows contact possibilities. *Infocom06* is a quite dense dataset with a high amount of nodes



Infocom05	contact	2	3	4	5	6+
average	18.3	19.5	6.8	3	1.3	0.5
maximum	101	215	87	46	29	23

Fig. 4. In *Infocom05*, distance 2 transmission opportunities overcome direct contact ones. This phenomenon is emphasized in dense periods. In such dataset, end-to-end paths using one relay are omnipresent and should be leveraged on.

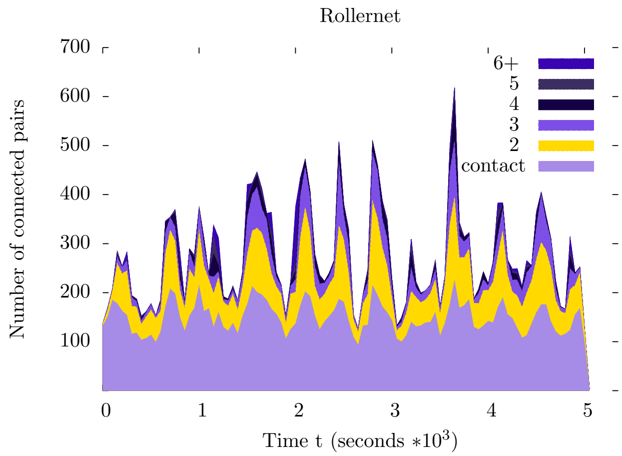


Infocom06	contact	2	3	4	5	6+
average	17.9	19.1	9.48	3.16	1.34	0.81
maximum	147	528	179	91	67	56

Fig. 5. *Infocom06* verifies the same properties as *Infocom05*. Distance 2 connectivity embodies more end-to-end transmission possibilities than mere contact.

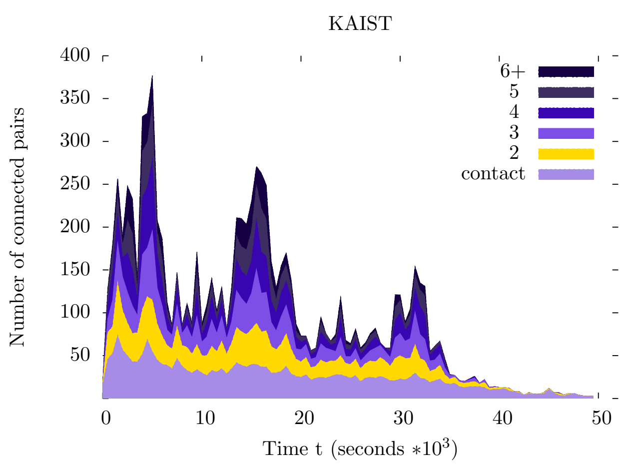
at regular pedestrian speed in a smaller surface than for other datasets we considered. This suggests that such environments, similar to many urban setting, are prone to have impressive extended end-to-end communication possibilities.

In Fig. 6, we observe *Rollernet* inherent accordion phenomenon [10] i.e., the sequential stretching and shrinking of the crowd due to urban obstacles preventing the crowd from moving forward. *Rollernet* has a dynamic setting with a compulsory path. Nodes do not have as much movement liberty as they have in other datasets. Contacts are prominent in *Rollernet*, however Favorable intercontact transmission are



Rollernet	contact	2	3	4	5	6+
average	147.3	76.6	32.5	9.8	3.1	4
maximum	250	245	182	83	37	6

Fig. 6. *Rollernet* being a specific sport event with a tight population, most transmission opportunities come from contact. However, during the revolving phase of the accordion phenomenon (density peaks), we observe a growth of distance 2+ paths.



KAIST	contact	2	3	4	5	6+
average	25.57	**19.75**	**17.72**	**15.4**	**11.05**	7.37
maximum	74	62	81	84	62	41

Fig. 7. *KAIST* is a campus wide experiment. Students are likely to spread around alleys and buildings giving a somewhat dense area in terms of taken paths. Therefore, we have similar Favorable intercontact proportion from distance 2 to 5. Opportunistic strategies clearly have room to spread in such setting.

still noticeable in peaks of density.

Considering *KAIST*, contact opportunities may seem more important than Favorable intercontact states. But, unlike previous datasets where connectivity quickly decreases with distance on average, we discover comparable connectivity between nodes for distance 2 to 5. *KAIST* environment being a campus-based measurement, it has a sparser density with a lot of movement freedom. Students tend to stay where other students are, like in restaurants, buildings, libraries but they

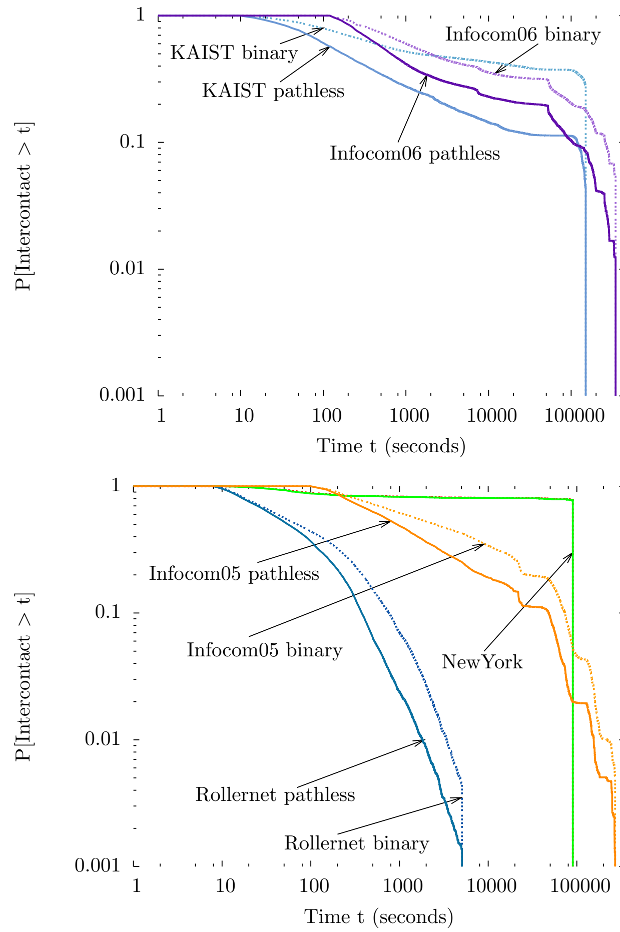


Fig. 8. Aggregated distribution of binary and Pathless intercontact times. Pathless intercontact keeps the same overall distribution as binary intercontact with a small deviation on the bottom left. In the *NewYork* dataset, both distributions line up as we have almost no Favorable intercontact.

may not share stronger relationships than just being student in the same university so they do not form close groups. This could explain the extended Favorable intercontact states here.

In any case, peaks indicate denser areas with higher communications possibilities. We visually notice how contact opportunities only represent a minor part of end-to-end transmission chances. Not considering Favorable intercontact ends up in another major network information loss.

Until now, we have focused on Favorable intercontact which is interesting because of its native communication opportunities. However, Pathless intercontact also brings interesting network indications. We will next focus on this aspect.

C. Pathless intercontact time characterization

Chaintreau et al. performed intercontact time characterization for various existing datasets [?]. They used the traditional binary contact vision where nodes not in contact are in intercontact. We perform the same analysis but with Pathless intercontact (time intervals deprived of contact or end-to-end paths between nodes) and compare distributions.

In Fig. 8 we plotted intercontact times CCDF i.e., the probability for an intercontact interval to last more than t seconds. Dotted lines represent the distribution for binary intercontact vision whereas solid ones correspond to our Pathless intercontact.

Distributions overall aspect remain the same with a general deviation to the left for Favorable intercontact CCDFs. This phenomenon is logical as we extended the contact notion, it reduces Pathless intercontact duration compared to initial binary intercontact. The contact extension only occurs in periods where the network is dense. As a result, Pathless intercontact distribution is not an exact translation of binary intercontact one but keeps an overall similarity. This also means that our definition does not change the intercontact properties our community used to have when studying DTN patterns. So, we can still leverage on the existing forwarding strategies background like techniques based on power law opportunities with adjusted parameters [9].

Our definition does not discard all the existing literature on DTN schemes, it enhances it by providing new beneficial network information via extended contacts.

V. Next Steps

Hybrid protocols. The inherent MANET nature of DTN should be taken into account not only for DTN understanding but also for their routing techniques design. Several inspiring routing approaches based their algorithm on this observation [11], [12]. Given our results, we suggest to pursue our research on hybrid protocols. They could mix DTN and MANET advantages: maintaining a close end-to-end vision around a node thanks to its immediate neighborhood and sending messages via modified DTN approaches for out of reach destinations. Via simulations, we next plan to implement and test such approaches.

Scope tuning. In highly dense areas, we obtain long end-to-end path but they may not be relevant for communication opportunities. In the case of routing protocols, an approach could choose to study favorable intercontact only up to a limited distance. The higher the n, the better the network comprehension but the costlier the computation. We can refine our n-ary intercontact definition with this regard in our future work on hybrid protocols.

VI. Summary and outlook

In this paper, we questioned the binary intercontact vision where nodes are either in contact else in mere intercontact. We advocate for a new ternary vision in DTN characterization via a fine-grained intercontact definition. Our contribution, the n-ary intercontact, proposes to enhance the intercontact notion. We expose two intuitive concepts: Favorable and Pathless intercontact. On a node-centered vision, Favorable intercontact brings new communication opportunities with the awareness of end-to-end paths. Whereas, Pathless intercontact includes the previous binary intercontact abstraction and preserves any existing approaches relying on its distribution.

Our n-ary intercontact characterization is very compliant to DTN applications. It grasps what binary intercontact vision misses: the underlying MANET nature in DTN. It explicits new communication opportunities. Here, we intentionally introduced a simple definition for our community to integrate it in their every day approaches. However, we already envision improvements concerning the routing research field.

We have shown that in *Rollernet* and *Infocom05* more than 50% of their participants display favorable intercontact opportunities. Maintaining the binary intercontact concept we miss all these end-to-end transmission chances. We have also seen how our Favorable intercontact transmissions exceed basic direct transmission but still maintaining low delay communications. So considering distance 2 end-to-end paths may double a node's chances to share its content with close by neighbors using low delay transmissions. This type of consideration adds new properties to DTN. They can guide new forwarding algorithms techniques like hybrid protocols we suggested.

We hope this work will bring incentives for our community to change its point of view on DTN characterization. We think the binary intercontact hypothesis is a major flaw and prevents disruption tolerant approaches from achieving their optimal efficiency when more information about the network can only help. In a near future, we plan on proposing hybrid protocols considering MANET and DTN transmission opportunities alongside to leverage communication opportunities revealed by the n-ary intercontact.

References

[1] K. Fall, "A delay-tolerant network architecture for challenged internets," in *ACM SIGCOMM*, 2003.

[2] P. Hui, J. Crowcroft, and E. Yoneki, "Bubble rap: social-based forwarding in delay tolerant networks," in *ACM Mobihoc*, New York, NY, USA, 2008.

[3] A. Lindgren, A. Doria, and O. Schelén, "Probabilistic routing in intermittently connected networks," *SIGMOBILE Mob. Comput. Commun. Rev.*, vol. 7, July 2003.

[4] T. Spyropoulos, K. Psounis, and C. S. Raghavendra, "Spray and wait: an efficient routing scheme for intermittently connected mobile networks," in *ACM SIGCOMM workshop on Delay-tolerant networking*, 2005.

[5] J. Whitbeck and V. Conan, "Hymad: Hybrid dtn-manet routing for dense and highly dynamic wireless networks," *Computer Communications*, vol. 33, August 2010.

[6] M. Musolesi and C. Mascolo, "Mobility Models for Systems Evaluation. A Survey," in *Middleware for Network Eccentric and Mobile Applications*, February 2009, pp. 43–62.

[7] S. Gaito, E. Pagani, and G. P. Rossi, "Opportunistic forwarding in workplaces," in *ACM SIGCOMM workshop on Online Social Networks*, 2009.

[8] V. Borrel, M. H. Ammar, and E. W. Zegura, "Understanding the wireless and mobile network space: a routing-centered classification," in *ACM MobiCom workshop on Challenged Networks*.

[9] A. Chaintreau, P. Hui, J. Crowcroft, C. Diot, R. Gass, and J. Scott, "Impact of human mobility on opportunistic forwarding algorithms," *TMC*, vol. 6, no. 6, pp. 606–620, 2007.

[10] P. U. Tournoux, J. Leguay, F. Benbadis, V. Conan, M. Dias de Amorim, and J. Whitbeck, "The accordion phenomenon: Analysis, characterization, and impact on dtn routing," in *IEEE Infocom*.

[11] J. Ott, D. Kutscher, and C. Dwertmann, "Integrating dtn and manet routing," in *ACM SIGCOMM workshop on Challenged Networks*, 2006.

[12] N. Sarafijanovic-Djukic, M. Pidrkowski, and M. Grossglauser, "Island hopping: Efficient mobility-assisted forwarding in partitioned networks," in *IEEE SECON*, 2006.