

Article

Corpus for Development of Routing Algorithms in Opportunistic Networks

Diego Freire , Carlos Borrego  and Sergi Robles 

Department of Information and Communications Engineering, Universitat Autònoma de Barcelona, 08193 Bellaterra, Spain

* Correspondence: diego.freire@uab.cat

Abstract: We have designed a collection of scenarios, a corpus, for its use in the study and development of routing algorithms for opportunistic networks. To obtain these scenarios, we have followed a methodology based on characterizing the space and choosing the best exemplary items in such a way that the corpus as a whole was representative of all possible scenarios. Until now, research in this area was using some sets of non-standard network traces that made it difficult to evaluate algorithms and perform fair comparisons between them. These developments were hard to assess in an objective way, and were prone to introduce unintentional biases that directly affected the quality of the research. Our contribution is more than a collection of scenarios; our corpus provides a fine collection of network behaviors that suit the development of routing algorithms, specifically in evaluating and comparing them. If the scientific community embraces this corpus, the community will have a global-agreed methodology where the validity of results would not be limited to specific scenarios or network conditions, thus avoiding self-produced evaluation setups, availability problems and selection bias, and saving time. New research in the area will be able to validate the routing algorithms already published. It will also be possible to identify the scenarios better suit specific purposes, and results will be easily verified. The corpus is available free to download and use.

Keywords: opportunistic networks; corpus; routing algorithms; scenarios; new communication paradigms



Citation: Freire, D.; Borrego, C.; Robles, S. Corpus for Development of Routing Algorithms in Opportunistic Networks. *Appl. Sci.* **2022**, *12*, 9240. <https://doi.org/10.3390/app12189240>

Academic Editors: Yang Yue, Runzhou Zhang, Hao Feng, Zheda Li, Lin Zhang and Dawei Ying

Received: 29 July 2022

Accepted: 13 September 2022

Published: 15 September 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

During the last decade, there has been some emerging networking paradigms that were announced that they will become mainstream in the future, such as Delay Tolerant Networking (DTN) or Opportunistic Networking (OppNet). Many of the use cases for them, though, are nowadays better solved by other approaches mainly based on global connectivity. Some examples of these use cases were providing connectivity in sparsely inhabited areas, in underdeveloped regions, and during disasters. However, there are still some scenarios for which flexible ad hoc communications without infrastructure require a different approach, closer to OppNets. Times have changed, and the overuse of the terms DTN and OppNet everywhere for many years has led to a current situation where these paradigms are regarded with suspicion by the research community, and even cause one to be wary of them. And yet, the need for this type of communications is still present. Perhaps it is more convenient to talk about concepts such as disruption-tolerant MANETs, multihop device to device routing in 5G, pervasive IoT, or dynamic source routing, but at the end of the day the concept of devices directly communicating with each other asynchronously using other peer devices as relays is still relevant to this day. Scenarios with these needs are, for instance, proximity-based applications, privacy-preserving communications, limited energy distance communications, or long-distance space communications, among others. For the rest of this study, we use the term OppNet to refer to this paradigm, regardless of the specific technology used to implement it.

Research on routing algorithms for OppNet is very important because it provides the core element that makes this technology work. Due to the asynchronous nature of the forwarding process, selecting the right neighbor to pass messages to, choosing a good number of copies of messages, and defining for how long messages are going to live in the network are crucial aspects of opportunistic networking that are included in the routing algorithm. High-performance algorithms are produced after following a rigorous process based on the scientific method, where evaluation, comparison, and testing allow to determine the best solution for a given scenario.

Unfortunately, most studies on this topic use an experimentation based on self-produced network traces, traces obtained in a particular real scenario by other researchers that have been adapted, or traces that were captured on a real network with specific constraints. Validity of results is often limited to a specific scenario or network conditions, and therefore these outcomes are often not an actual indication of the universal validity of the solution that would allow its utilization by the global community. Obviously, this is not what researchers want, and there is no hint of any bad intention here. The reason of this situation is the lack of common frameworks and objective test environments for facilitating the production of quality routing algorithms enabling the optimal application in the scenarios requiring them.

To solve this problem, methodological approach is required that is based on scientific rigor. Being able to design good algorithms that do not just hold water but that can be objectively evaluated and compared to others under fair conditions is the only way to choose the best option for a particular scenario.

In this study, we propose the cornerstone of this methodology: A corpus of carefully selected scenarios, accessible to the entire community, that can be used for the actual comparison and assessment of routing algorithms. This has not been an easy task. It might seem that just selecting some common, already published scenarios would suffice for this objective. However, that would be incomplete. A valid corpus has to be representative of all possible scenarios; it has to be accessible to the entire community, using a standard format; and all of its scenarios have to be comprehensive, with no missing parts that could be completed in different ways. Similar initiatives are found in other domains, such as image compression. We have studied the different variables, the dimensions, and defining a scenario, finding out a subset that can be considered independent to form a sound vector base for the scenario space. Then, we have selected forty-one of them to constitute the corpus. We have tested this corpus by using a high replication algorithm, and have observed that it performs differently for all of the scenarios.

This corpus is a leg to stand on for new research on the area. At last, fair comparison can be done, and results are easily verifiable. It can also validate the already published routing algorithms and help determine the scenarios they suit better.

The rest of the article is structured as follows: Section 2 describes all of the relevant state-of-the-art information, paying special attention to routing algorithms in opportunistic networks, the current evaluation approach that routing algorithms have, and a review of how other research fields conduct the evaluation of algorithms. Then, in Section 3, the article provides a complete description of a new methodology for evaluating OppNet routing algorithms. The article follows with Section 4, where the appraisal of the contribution is presented with a simulation-based experiment. Next, Section 5 contains a discussion, and finally, Section 6 presents the conclusions drawn from this work.

2. State of the Art

In this section, the state-of-the-art of opportunistic networks is reviewed, emphasizing the performance evaluation of routing algorithms. Then, this article provides an overview of the tools, strategies and metrics used to evaluate the performance of routing algorithms. Additionally, this section describes the challenges when evaluating and comparing opportunistic networks. Finally, we review how other fields have tackled similar problems to assess the performance of algorithms.

2.1. Opportunistic Networks

Opportunistic network(s) (OppNet(s)) are wirelessly connected devices that interchange information, exploiting connection opportunities. In this type of networks, devices with wireless capabilities (such as smartphones, tablets and smartwatches, among others) use direct communications opportunities [1]. OppNets allow information exchange among devices even when an end-to-end path may never exist [2]. Moreover, the variations in the network's topology are considered normal behavior due to the wireless nature of the devices [3].

Additionally, OppNets are challenging networks where disruption and delays in communication are considered normal [4], due to tackling the problem of how to exchange information without a fixed network infrastructure [5]. In OppNets, the information transmitted between devices is also known as messages. These types of networks use the store-carry-and-forward paradigm to transmit information among devices [6,7]. This paradigm allows message routing from source to destination, handling disconnection, delay and disruption in the network. When a device implements the store-carry-and-forward paradigm, the device receives a message. Next, the message will be stored and carried until a transmission opportunity occurs, and finally, the message is forwarded to the other device. In OppNets, a device is also known as a node [8].

The applications of OppNets have been widely studied. In environments where traditional networks do not perform well or, even worse, cannot operate, OppNets may provide a feasible solution for communication. Among the challenging environments, research highlights the following as areas where OppNets may perform well: Cellular network offloading [9], communication in challenged areas [9], censorship circumvention [10], mobile ad hoc social networks [2], offline social networks [2], Internet of Vehicles [11], information-centric networking [12], and proximity-based applications [2], among others.

OppNets are an active research field that is still worth to be studied. For example, one interesting open research topic beyond OppNet applications that inherently implement a store-carry-and-forward delivery paradigm is information-centric networking (ICN) [13]. This communication architecture can effectively suit OppNets. ICN is a non-host-centric communication architecture that, unlike IP, is not tied to a specific network location. It is centered around hierarchical content names used directly at the network layer [14].

A routing algorithm can be described as an implementation of a message routing function whose objective is to deliver the messages to their destination while maximizing the efficiency of resource consumption. Routing algorithms are the intelligence that supports the operation of an OppNet, since they dictate the directives on the behavior of the nodes with the messages. Routing algorithms seek to maximize delivery by optimizing the use of resources [15]. Over the years, researchers have put their efforts into developing routing algorithms. Articles such as [16–18] mention a number of algorithms that have been proposed. These proposed routing algorithms provide routing solutions for particular environments. Some routing forwarding strategies implementations use epidemic [19], probabilistic [20], number of copies [21], or based on neighborhood contact history [22] strategies to deliver messages among nodes.

In the context of forwarding decisions, the routing algorithm must decide upon the best candidate(s) to receive a message among all available nodes. In addition to forwarding, a message has a lifetime in the network, and the routing algorithm will update the message lifetime. Messages can also be stored and deleted; this algorithm does not require any selection of peers.

So far, the definition of OppNets was presented and it was shown that routing algorithms are the basis of communications in challenging environments known as OppNets. The critical role routing algorithms play in an OppNet was also shown. The following section explains how OppNets and routing algorithms can be described.

2.2. Characteristics and Metrics in OppNets

In this section, the article explains the characteristics and metrics used in OppNets and why they are fundamental concepts in OppNets. This section also describes the relationship between the characteristics and the metrics.

OppNets are heterogeneous. A specific OppNet instance can be expressed with a set of characteristics. However, it cannot be said a priori that one instance of an OppNet is necessarily equal to or different from another OppNet. One way to establish the differences between OppNets is to compare the characteristics that describe each OppNet.

The more characteristics are used, the more accurate is the description of an OppNet's behavior. In other words, this feature-based description somewhat simplifies a real-world OppNet. As said in [23], characteristics are deployment facts expressed in numbers for a network.

Metrics return quantitative information about a feature or behavior. Moreover, a metric is a measure function whose output is a numerical value that can be interpreted as the degree to which the routing algorithm has a given attribute [24]. Researchers use metrics to evaluate, compare, or measure the behavior of OppNet routing algorithms [25,26]. Metrics quantify, among others, the performance and the ratio that routing algorithms achieve when the messages are interchanged between source and destination. It can quantify a specific attribute (such as a count of successful processes, time consumption, and messages delivered), providing a quantitative indication.

There are some metrics that most OppNet researchers tend to use to prove performance hypotheses. Among them, three stand out because of their presence in most works related to OppNets this work has found. Those metrics are: Delivery ratio, delivery delay and delivery cost [16]. However, some authors do not use these metrics but rather modified versions of them to fit specific hypotheses. In other cases, some authors even find it necessary to establish entirely new metrics to measure the behavior of their work [27,28].

Characteristics and metrics have a close relationship. This relationship is given because routing algorithms work in OppNets instances. Moreover, since OppNet instances are described by a set of characteristics and those characteristics, in a way, produce a routing algorithm behavior measured by the metrics, the characteristics of an OppNet influence the metrics that a routing algorithm has. For example, there are equivalent network configurations where the same message routing algorithm would be equally efficient. However, the metrics will tell if two scenarios are the same or different from a routing point of view. An adequate characterization of an OppNet withdraws the attention on details that can give a wrong cognitive impression of the difference between networks.

This section has described the characteristics and metrics in OppNets, and it also has described the relationship between them. The following section shows the current performance evaluation techniques among OppNet routing algorithms.

2.3. OppNet Routing Algorithms Evaluation and Comparison

Current evaluation and comparison techniques are worthy of being explained. This section explains the evaluation and the comparison of routing algorithms. Furthermore, it is described the one-way connection between the evaluation and comparison of routing algorithms in OppNets.

Transforming a routing idea into a routing algorithm is a challenge by itself. A complete creation methodology helps in that matter, increasing the quality and speeding up the creation process [29]. Previous work by the authors [29] showed a seven-stage methodology for developing new routing algorithms. This methodology is depicted in Figure 1. The first stage of the seven-stage methodology is the routing idea, where the concept of the routing mechanism is conceived. In the second stage, the idea is modeled; thus, in the third stage, the routing proposal can be analyzed. After the conception, modeling and analysis, the fourth stage simulates the routing algorithm. However, a successful simulation does not guarantee real-world implementation. The fifth stage requires a full-featured code capable of featuring in the real world. In the sixth stage, real-

code is executed in controlled conditions—often a proof-of-concept. Finally, the application phase is where the routing algorithm is deployed in a real-world environment with real devices and users. If the creation of a routing algorithm follows a creation methodology, results can be evaluated, compared and repeated. A reliable routing algorithm design methodology enables an objective method of evaluating and comparing routing algorithms.

From the previous paragraph, the reader may note that evaluation and comparison are different terms. Evaluating a routing algorithm can be described as an intra-technique [30] that quantitatively recognizes the routing algorithm behavior for a particular OppNet environment. The evaluation process of one does not require other routing algorithms to assess their performance.

On the other hand, the comparison among routing algorithms can be described as an inter-technique [30] that ranks the performance of a routing algorithm against other routing algorithms.

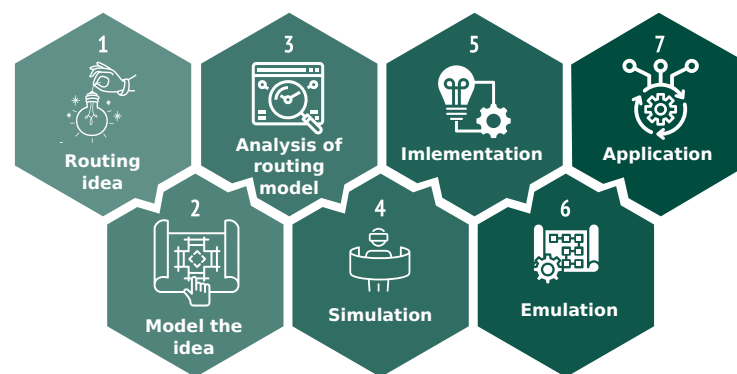


Figure 1. Seven-stage methodology for routing algorithm creation [29].

Although the evaluation and comparison of routing algorithms are closely related, some differences exist. A comparison of routing algorithms requires an evaluation of the performance of any individual routing algorithm, which means that comparison is not possible without the evaluation. However, some authors might be interested only in the evaluation rather than the comparison.

In general terms, metrics can be used to evaluate and compare the performance of routing algorithms. In OppNets, metrics can be obtained from simulation tools. However, comparing the metrics of different algorithms does not address a fair comparison by itself. A fair comparison among routing algorithms could be assessed when OppNets environments, messages and simulation settings are equal or equivalent among them. This section explained the evaluation and the comparison of routing algorithms. Furthermore, it also described the one-way connection between the evaluation and comparison of routing algorithms in OppNets. The following will explain how OppNets have been simulated nowadays.

2.4. OppNet Simulation Deployment Nowadays

This section presents the main existing OppNets simulation tools. It also illustrates the elements and parameters that allow an OppNet simulation deployment and where those tools come from. Furthermore, this section reflects upon how current simulation tools are used.

Figure 2 shows the elements involved when simulating an OppNet. These simulation elements are input, output and software setup. The inputs define the behavior of the network, for example, nodes and message characteristics. Instead, the simulation's output is the information obtained after the simulation, for example, routing algorithm performance metrics and delivery information of messages. Most of the time, performing a post-simulation analysis from the data obtained as the output may be necessary. It is expected that different simulation parameters return different outputs since the simulation is sensible to setup changes [31].

An OppNet has several software simulation alternatives. Among the software tools that allow simulating an OppNet are GloMoSim [32], OMNeT++ [33], DTN2 [34], HuggleSim [35], the ONE Simulator [36], ns-3 [37], Adyton [38] and MobEmu [39]. The simulation tools are mentioned in ordered of creation from 1998 to 2018. For OppNets research, the most used is the ONE simulator, reaching 62% of recent publications [16].

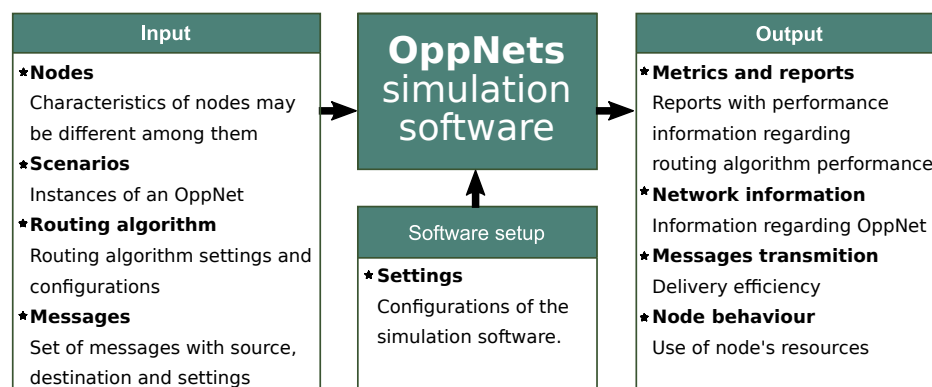


Figure 2. Scenario input, output and configuration elements that enable an OppNet simulation.

Furthermore, as is shown in Figure 2, scenarios are inputs of a simulation. The literature uses the word scenarios interchangeably with the elements they refer to. That is, there is no definition of a scenario in OppNets, and most authors refer to the contact traces (also known as traces or mobility datasets) as scenarios. According to [40], traces are datasets containing registers of nodes, and the information is either positions, contacts or both during a time. Some OppNet simulation softwares, such as the ONE simulator [36] accept trace datasets as input.

Concretely, the source of these traces can be real-world, synthetic or hybrid. The synthetic traces can be produced far faster than real-world traces and may be as valuable as their real-world counterparts for evaluation purposes [41]. A hybrid trace is a mix of real-world and synthetic traces; there is no predefined portion of the real and synthetic traces.

However, the non-standard real-world traces have a cognitive bias. Although they may have some realistic characteristics, their random nature makes generalization difficult and may not be suitable for different environments. For example, if two connectivity traces have been collected from two universities, their characterization will be similar and may not be suitable for simulating a countryside OppNet.

On the other hand, and in addition to the non-standard traces, the synthetic traces are a feasible solution for representativeness because by having control of the characteristics they represent, it is possible to select the traces that, as a whole, are a better representation of a desired OppNet environment. Moreover, since the interest is to represent the real world, it is better to have traces that, due to their characteristics, are representative of the real world instead of several real traces that have similar network behaviors, regardless of the origin. The importance of a trace stands on its network behavior rather than the creation origin.

In the same way, there are several traces suitable for OppNets. Sites such as CRAW-DAD [40] gather mobility traces datasets that are shared among the scientific community. Indeed, CRAW-DAD has 135 datasets (reviewed on 15 June 2022), but despite this amount, a few datasets are often used rather than others. Some studies even call those “well-known traces” [42] or “well-known scenarios” [43,44]. Datasets such as Asturias [45], Taxis Roma [46], Taxis San Francisco [47] and Cambridge/Haggle [48] are some of those that are usually included in the literature as “well-known traces”.

Authors interested in the evaluation and comparison of routing algorithms that might use datasets such as Asturias [45], Taxis Roma [46], Taxis San Francisco [47] and Cambridge/Haggle [48] should have complete knowledge of the representativeness of those datasets. However, evaluation and comparison are not a matter of the number of traces

instead of network behaviors. It should be the focus of the authors to test routing performance in representative scenarios. For example, in terms of network behavior, traces do not include information to state a difference between the non-standard traces of Taxis Roma [46] and Taxis San Francisco [47], because both represent the mobility of taxis in a city.

A common practice to assess routing performance is the comparison against peers. Until now, the routing performance has been assessed based on how better an algorithm is compared to other selected routing algorithms within a set of non-standard network traces. However, a routing performance evaluation cannot be extended outside the specific routing algorithms and non-standard network traces. Nowadays, literature does not have a benchmarking scheme for the performance of routing algorithms [16].

In this section, the reader could have seen how the scientific community naturally looks for a group of traces that, in some way, standardize the environments to evaluate the routing algorithms. The following section reviews other research fields with similar problems of evaluating and comparing algorithms, the approach and, above all, the solutions they have found, even though the algorithms described in the following section are not routing-related.

2.5. Algorithm Performance Evaluation in Other Fields

The previous sections introduced the features, behaviors and characteristics of Opp-Nets and the importance of routing algorithms. It also showed pitfalls for a fair comparison among routing algorithms. The following section reviews how other fields have proposed solutions for fair comparisons. Specifically, this section reviews how the fields of data compression, linguistics and speech recognition handle the performance comparison problem when developing new algorithms.

This section introduces the term corpus, which refers to a collection of representative data used to analyze the effectiveness of an algorithm's behavior.

2.5.1. Data Compression

Data compression aims to reduce the volume of data while preserving the quality, and it can be classified as either lossy or lossless compression. In lossy and lossless compression, the goal is to maintain quality by using the least amount of data to represent the information. In lossless data compression, the original data can be obtained. However, in lossy compression, some information is lost.

A corpus, in data compression, is a collection of representative files to evaluate the effectiveness of the compression ratio [49]. Calgary [50] and Canterbury [51] are corpuses used in lossless data compression.

Using a corpus to evaluate compression algorithms reduces bias and facilitates the experiments' reproducibility. Furthermore, using a corpus creates compression benchmarks, a standard compression ratio that other algorithms may be compared to. Nowadays, the criteria regarding the corpus are widely accepted in the compression field.

2.5.2. Linguistics Corpuses

As in the field of data compression, linguistics corpuses are sets of text used to study language composition. The use of a corpus allows, in the case of the field of linguistics, to extract complex language structures, which could not be extracted without having a collection that has these complex language structures represented in its files.

Using a corpus can broaden research in other fields. In the case of linguistics, dictionaries and translations have benefited from using a corpus.

2.5.3. Speech Recognition

In speech recognition, the use of corpuses when comparing results is extensive and diverse. The number of corpuses results from one language's heterogeneity compared to another. In languages it is complex to recognize speech, because one language's accents

differ from other dialects and phonetics. Nevertheless, despite the variety of results, in this research field, a corpus is a set of selected files seeking representativeness, limiting the number of elements to those necessary, widely available and valuable for developing and evaluating speech recognition techniques.

The techniques of compression, speech recognition or routing will be useless if applied to data that are not relevant or representative. A corpus is helpful within the intended scope of usability.

In this section, our research showed that the creation of routing algorithms voids a fair comparison. Furthermore, this section showed that using self-selected files to perform compression algorithm comparisons seems similar to the well-known traces used to compare routing algorithms. The insights obtained from the review of algorithm evaluation and comparison are that using a corpus improves the performance of algorithms throughout standardization.

3. A Corpus for Routing Evaluation in OppNets

Section 2 has shown that developing routing algorithms in OppNets can be improved using an algorithm creation methodology, particularly when comparing results. Although comparison is essential in research, scientific rigor cannot be assessed now when comparing the performance of routing algorithms. Section 2 also showed that a corpus helps in the algorithm development process, proving to be a crucial part of the methodology. This section defines what an OppNet scenario is and how it can be characterized. Next, this section defines a complete methodology for the development of a corpus. Finally, this section presents a corpus for evaluating and comparing routing algorithms.

3.1. Scenario Definition

As is explained in Section 2, nodes are the principal component of an OppNet scenario. Nowadays, the scenarios are considered a time-ordered list of contacts or positions that nodes have within the same OppNet. It is also mentioned in Section 2 that this information has been called contact traces. However, the contact traces also contain, in a non-explicit way, the corresponding network behavior. Characterizing a trace describes the intrinsic network behavior of the trace with a vector of characteristics. In this article, *an OppNet scenario is denoted as a trace of positions characterized by a vector of seventeen characteristics*.

3.2. Scenario Characterization

Characterizing a scenario basically consists of defining the characteristics that describe their network structure and behavior entirely. This study identifies two types of characteristics, namely direct and indirect. Direct characteristics are the ones that can be identified or defined directly, for example, by counting the number of nodes or measuring the speed of the nodes. On the other hand, indirect characteristics represent characteristics that can not be directly configured but, instead, can be estimated. For example, nodes' centrality refers to betweenness centrality in a trace and cannot be configured with state-of-the-art tools.

A trace could be considered a scenario after the trace has been characterized, that is, when the contact or position trace has a vector describing its network structure and behavior. The characteristics had been identified from the literature review of OppNets routing algorithms. Table 1 shows the list of seventeen characteristics that has been used to characterize the network behavior of a contact trace. Seven characteristics are direct (D) and the rest are indirect (I). The number of nodes, node speed, studied area, movement pattern, node centrality, node contact time, and total encounters are among the direct and indirect characteristics.

The measurement of the characteristics only concerns indirect characteristics. The measurement of those indirect characteristics listed in Table 1 follows the directives depicted in Table 2 and the next paragraph.

Table 1. Set of characteristics for a scenario definition, characteristics are classified as direct (D) and indirect (I).

N°	Characteristic	Type	Description
1	Total number of nodes	D	$[nodes] \Rightarrow \{nodes \mid 192 < nodes < 960\}$
2	Nodes per group	D	$[nodes_by_group] \Rightarrow \{2^n \in \mathbb{Z} \mid 3 < n < 10\}$
3	Groups of nodes	D	$\{groups \in [1, 2, 3, 4]\}$
4	Node's movements	D	$[movement] \Rightarrow \{movement \in [m_1, m_2, \dots, m_m]\}$
5	Node's speed	D	1, 3, 7, 14 and 27 m over second
6	World size	D	$[width, height] \Rightarrow \{[width, height] \mid width, height \in [200 \dots 3200]\}m$
7	Area	D	$[area] \Rightarrow \{[area] \mid area \in [4000 \dots 4,160,000]\}square\ meters$
8	Centrality	I	Measure of how much a given node is in between other nodes
9	Inter-contact time	I	Time a node has no connection
10	Contact time	I	Duration time of the connection between two nodes
11	Contact time per minute	I	Contact time within a minute window
12	Contact node ratio	I	Ratio of nodes contacted by a node
13	Popularity	I	Measure of the ratio of total unique connections
14	Window centrality	I	Mean centrality in a period
15	Encounters	I	Number of encounters
16	Sociability	I	Ratio of contacts
17	Total encounters	I	Total number of encounters within nodes

In Table 2, the number assigned to the characteristic corresponds to the number defined in Table 1. With the exception of characteristic number seventeen (total encounters) the characteristics shown in Table 2 are calculated in a two-step process. The first step is to calculate the characteristic individually in each node. Then, as a second and final step, the mean, variance and standard deviation values of the characteristics are calculated within the values of all or some nodes included in the scenario. For characteristic seventeen, the second step is the sum of the individual values of all nodes. The particular considerations are listed as follows:

- Centrality, inter-contact time, contact time, contact node ratio and encounters: mean of the individual measures of all nodes.
- Popularity and sociability: mean of highest ten percent measurements.
- Contact time per minute, window betweenness centrality: mean of metrics within a period.
- Total encounters: accumulative measurement.

The following section introduces the concept of a corpus in the context of the OppNets. The concept of a corpus will be used across this article as the principle for standardizing the evaluation and comparison of OppNet routing algorithms.

3.3. Corpus Definition

A corpus, in the context of OppNets, is a collection of OppNet scenarios with two main features: First, all scenarios work together to cover all possible network behaviors,

and second, the routing algorithms have different performance behaviors when routing messages in each scenario.

Table 2. Indirect scenario characteristics measurement directives with references.

N°	Characteristic	Measurement Directive	Ref.
8	Centrality	Betweenness centrality computed as number of connections held by each node	[2]
9	Inter-contact time	Elapsed time each node has between contacts	[52]
10	Contact time	Elapsed time of the connection between two nodes	[52]
11	Contact time per minute	Contact-time within a period of one minute	[53]
12	Contact node ratio	Node contact ratio	[54]
13	Popularity	Unique peer-connections a node has	[55]
14	Window centrality	Centrality during a period	[56]
15	Encounters	Number of connections a node has	[57]
16	Sociability	Ratio of the number of contacts a node has to the total number of nodes	[58]
17	Total encounters	Summation of the number of connections within nodes	[57]

3.4. Quality Requirements

The corpus aims to be a fair field for evaluating OppNet routing algorithms, providing a set of scenarios that can emulate real-world environments due to their characteristics. This article presented a corpus creation methodology depicted in Figure 3 and explained it in detail in Section 3.5. In addition, the corpus creation methodology presented in this research pursues the following requirements: Coverage, scope, quality and usability.

- Coverage: the coverage of the corpus should have representativeness for real-world environments, considering a significant difference between scenarios.
- Scope: the scope of the corpus should be the performance evaluation of routing algorithms in OppNets.
- Quality: the quality of each scenario of the corpus should be guaranteed by analyzing the representativeness and diversity among other scenarios.
- Usability: the corpus should be easy to use, and the scenarios should be adaptable to simulation software, where the evaluation of the performance of algorithms in OppNets is carried out.

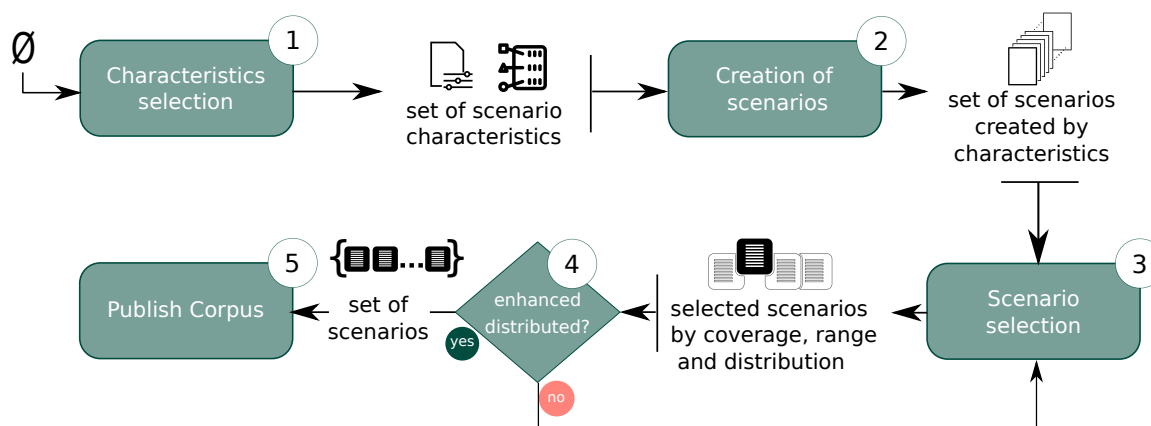


Figure 3. Corpus creation methodology with backtracking stage for scenario selection assuring purpose, coverage, scope, quality and usability requirements.

3.5. Corpus Creation Methodology

This section presents the corpus creation methodology depicted in Figure 3. The corpus creation methodology has five well-delimited stages, each with specific inputs, outputs and tasks. The input information of one stage is the output of the previous one, except for the first stage, which does not have an earlier stage.

The first stage, characteristics selection, decides those characteristics that describe a scenario. The selected seventeen characteristics are displayed in Table 1. A Pearson correlation [59] study of the selected characteristics was performed as shown in Figure 4. Some characteristics have a high correlation because they are based on connections and interaction between nodes. However, despite the redundancy and the high correlation, the characteristics reflect essential connectivity behaviors of the scenarios. This is why these highly correlated characteristics remain within the selected characteristics.

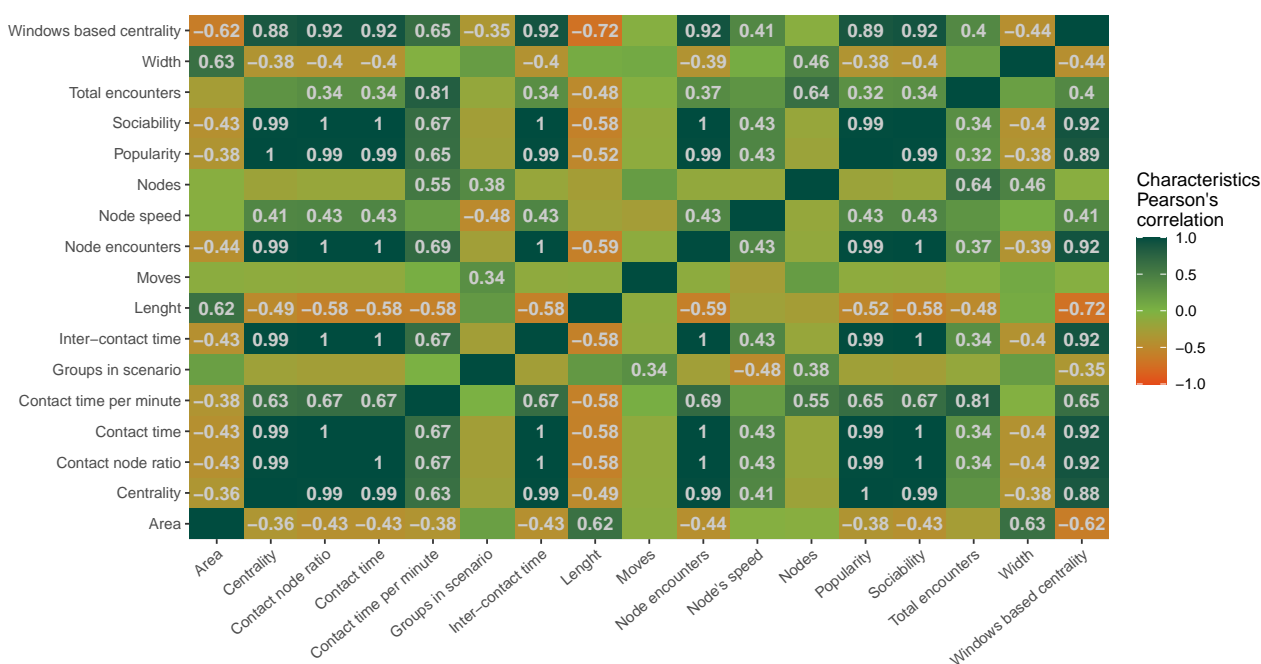


Figure 4. Heatmap of Pearson correlation coefficients between scenarios' characteristics. Only significant Pearson's correlation coefficients are shown.

Figure 4 made it clear that some characteristics were highly correlated. At first glance, one way to deal with highly correlated characteristics is by removing them. However, removing characteristics that describe scenarios was considered a wrong approach because fewer scenario characteristics might hinder the scenario description accuracy. The high correlation helped us understand that selecting scenarios would not be straightforward and that it will require a backtracking process to achieve diversity and representativeness among scenarios in the corpus. The backtracking process uses additional information about the characteristics. Specifically, the variance of the characteristics was used in the case there is a need to achieve representativeness and diversity objectives.

The second stage, creation of scenarios, received the characteristics found in stage one and then created scenarios for the given characteristics. This stage generated over 200,000 OppNet scenarios, many of which had similar behaviors and therefore similar vectors of characteristics. The scenarios with a similar vector of characteristics were considered equivalent.

The third and fourth stages, scenario selection and enhanced distribution, were loop-connected. Each characteristic range was evenly divided into sub-ranges called windows. Then, a subset of scenarios was selected for each window, and this process sequentially looped through the list of characteristics. The number of scenarios was reduced because the

scenarios should belong to all windows of the characteristics. If there was no scenario in the window, those empty-scenario windows were re-adjusted until scenarios were found.

When all of the characteristics had been run through, and a representative number of scenarios had been obtained, stage four checked the diversity of the scenario collection. The loop was broken if the diversity of scenarios was fulfilled, which implies not having similar scenarios and that the distribution of characteristics manages to cover the entire range of each characteristic. Each scenario fulfills a part of the range of the characteristic. All scenarios, as a whole, complete the range of the characteristics.

The final stage, publish corpus, made the corpus of OppNet scenarios available for the research community. This assures the usability set as a quality requirement shown in Section 3.4. The following section describes the corpus obtained following the corpus creation methodology presented in this section.

3.6. Corpus Morphology

Section 3.5 describes the creation of the corpus of OppNet scenarios that address the quality requirements mentioned in Section 3.4. Creating the corpus following the methodology returned forty-one scenarios with a balance between representativeness and diversity. The similarities among the scenarios increased with a number higher than forty-one, thus harming the diversity of the corpus. Moreover, some characteristics were not represented when the number was lower than forty-one. Therefore, the corpus is a collection of forty-one OppNet scenarios, and the characteristics and their distribution can be seen in Table 1 and Figure 5, respectively.

Scenarios in the corpus are identified with a number in the range [1–41]. Additionally, the corpus covers the range of each characteristic with the range of each scenario. In Figure 5, the X axis of each sub-figure represents the scenarios, and the Y axis represents the characteristic. Scenarios depicted in Figure 5 are not ordered by their number but by the value of the characteristic.

Furthermore, Figure 5 shows that node centrality, node inter-contact time and node sociability are characteristics with a high Pearson correlation among them. That is the reason why their figures have resemblance among them.

Figure 6 shows a study of the diversity of the corpus scenarios using a heatmap. It shows the relative intensity of characteristics of each of the scenarios in the corpus. Each column in Figure 6 is a scenario of the corpus. As it is mentioned throughout this article, the corpus will be expected to have representative as well as diverse scenarios. Figure 6 shows that (1) there are no equal scenarios and (2) the distribution of the characteristics is uniform since there is no predominance of a single color.

For usability reasons, each scenario of the corpus has two types of traces mentioned in Section 2.4, the contact traces and their homologous position traces. Furthermore, the granularity of the position traces is one second. Additionally, the contact traces can be obtained from their homologous based on the node positions but not the other way around.

The scenarios simulate the speed of pedestrians, cyclists and two types of motorized vehicles. Those speeds are shown in Table 1. For this reason, up to four groups have been organized for each stage. Nodes among the same group share the speed and movement pattern. Movement patterns and node speeds are described in Table 1.

The number of nodes present in a scenario differs from one scenario to another. Still, the total number of nodes is distributed unevenly among the groups present in the scenarios with more than one group.

The morphology of the corpus depicted in Figure 5 is well distributed as a result of the methodologically selected scenarios. The following section assesses corpus behavior when routing messages with routing algorithms.

Section 3 explained the concept and characterization of an OppNet scenario. It also defined and created a corpus to evaluate and compare OppNet routing algorithms. Section 3 also described the creation methodology and the morphology of the corpus obtained. The following section assesses the behavior of the corpus.

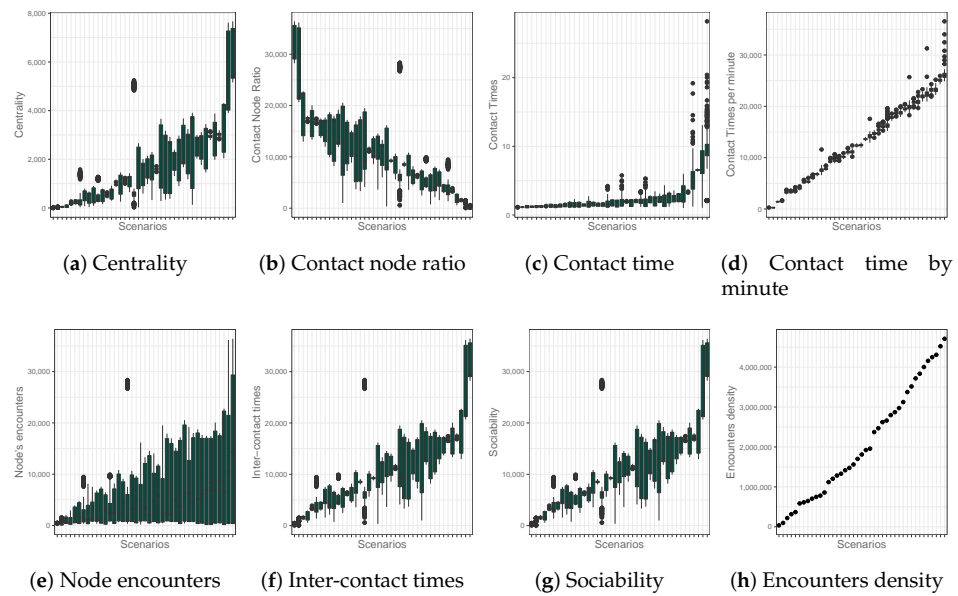


Figure 5. Scenario characteristics range distribution.

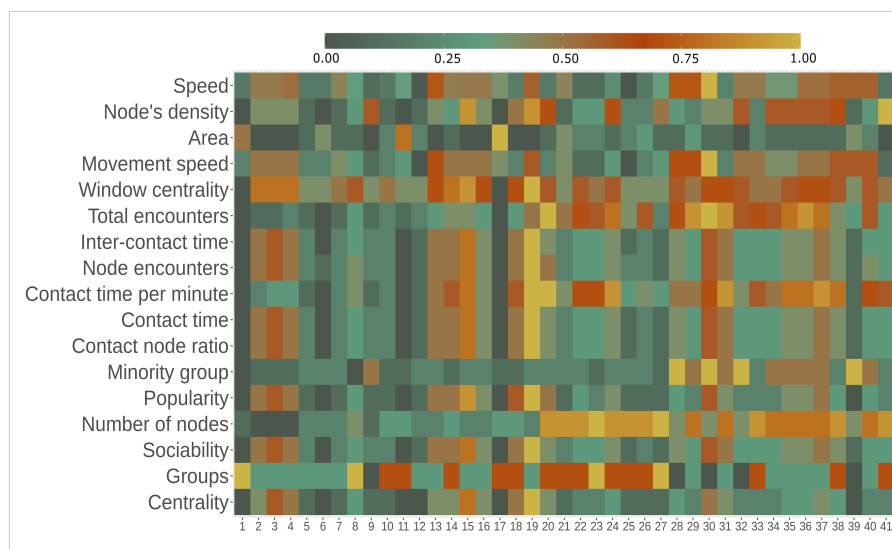


Figure 6. Diversity representation of the forty-one scenario collection, which constitutes the first corpus for the performance evaluation of OppNet routing algorithms. The X-axis is the scenario's number, and the Y-axis represents the characteristics.

4. Corpus Appraisalment

In this section we aim to appraise the corpus behavior when routing messages using a concrete routing algorithm. For this purpose, a series of simulations were conducted to depict the behavior of the corpus scenarios. Therefore, a routing algorithm with high replication of messages was selected. The reason for such a selection is to verify if, under intense replication conditions, the corpus shows a different response within the scenarios.

4.1. Corpus Performance Appraisalment

The experiment was conducted over the opportunistic network environment (the ONE simulator) [36] using the corpus of OppNet scenarios presented in Section 3.

Forty-one simulations were performed to assess the network representativeness of the corpus. In those simulations, node and message configurations were equal for all simulation setups. The forty-one simulations mean one simulation for each scenario of the corpus. The

routing algorithm was an epidemic algorithm, a routing algorithm replicating messages to every contacted node. The reason behind the selection of an epidemic routing algorithm for the experiment was the ability to flood the network with messages exhaustively. An epidemic routing algorithm will forward a message to every node that it has contact with. Then, each recipient node will store the message until a new connection arises and repeat the forwarding process. An epidemic routing algorithm will delete the message only when the assigned time to live of the message is reached.

As was explained in Section 2, routing algorithms aim to transmit messages from source to destination. For this reason, network behavior could be expressed by how messages are delivered within the scenarios in the corpus. The simulations of the experiments have shown the behavior of the corpus with the metrics related to message delivery. The metrics analyzed were: The number of messages delivered, messages relayed, messages aborted, messages dropped, message hop-count and the message buffer time.

Figure 6 depicts the diversity within the characteristics vectors that define the scenarios in the corpus. In order to establish a difference among scenarios and, therefore, the corpus reliability, the scenario responses should be different between them. The response generated by each simulation was analyzed graphically to find their differences. Figure 7 presents the differences between the behaviors of the scenarios.

Figure 7 shows the differences of the response with eight sub-figures. Each sub-figure is a different metric. The *Scenarios* axis in each sub-figure stands for the forty-one scenarios. Although all sub-figures contain the same scenarios, scenarios are not ordered equally from one sub-figure to another because they are arranged in ascending order according to the metric that sub-figure represents. The Y axis in each sub-figure represents the normalized value of each scenario. Furthermore, each sub-figure depicts forty-one values in the [0–1] range since values are normalized.

The results show a different response from one scenario to another, proving a different behavior in each scenario. These results show the diversity among scenarios, which is expressed in Section 3.4 as a corpus design requirement. Some areas are denser than others, but responses are well distributed overall.

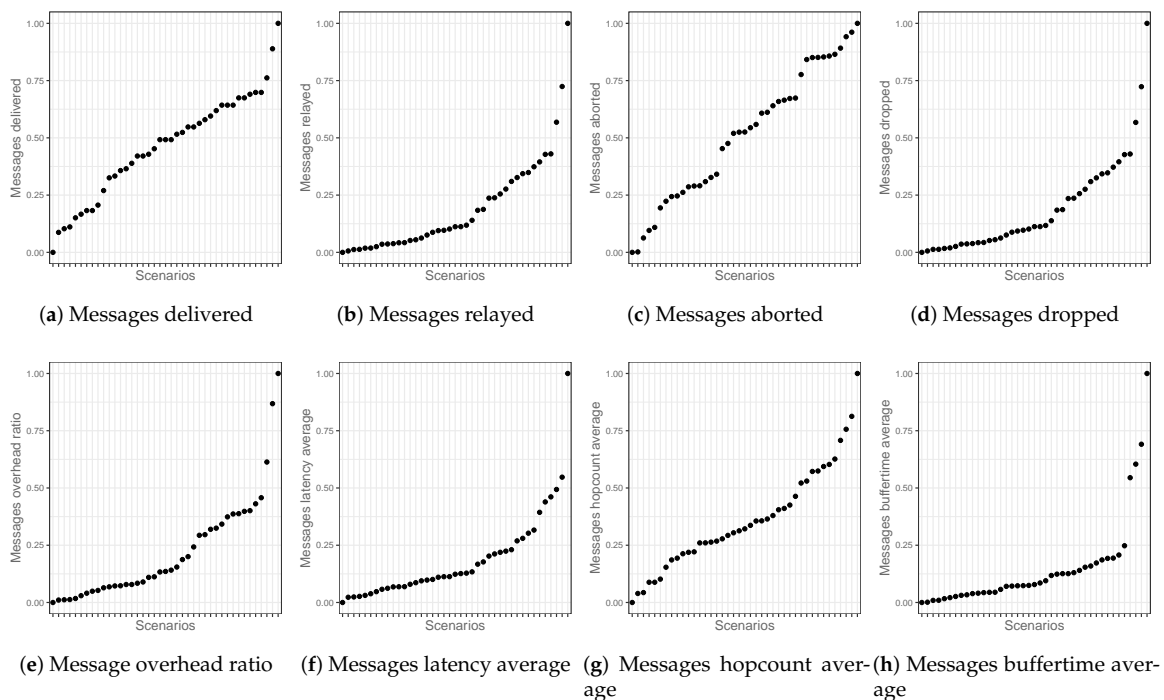


Figure 7. Corpus scaled benchmarks for epidemic routing. Each dot in the figure is a scaled outcome ordered by the the metric; the identification number of scenarios is not shown and the order changes among sub-figures.

4.2. Evaluation and Comparison Using the Corpus

Now that the corpus contribution has been obtained via the methodology shown in Figure 3 and explained in Section 3.5, this section describes how the corpus can be used when a routing algorithm's evaluation and comparison process is needed. For the sake of clarity, some in-depth details are not included in this section, such as software configurations. The reader is asked to keep in mind that this section is intended to outline the usability of this study's main proposal rather than providing a closed recipe for using the corpus of OppNet scenarios.

When the research stage requires an evaluation of a routing algorithm, researchers interested in using a corpus will have to implement a simulation environment such as the one shown in Figure 2. The researchers should start by configuring the OppNet simulation software. After this, to use the corpus, researchers will have to download it. The corpus is available entirely free and without the necessity of login information. It is understood that at this stage, the routing algorithm that is going to be evaluated is already selected. Finally, researchers might configure nodes and messages and establish the metrics that will retrieve the information to evaluate the performance. If the researcher desires to assess a comparison, the process will have to be repeated only by changing the routing algorithm. Then, the researchers should compare the corresponding routing metrics obtained from the respective simulations.

In this section we have evaluated the behavior of the corpus with a high replication algorithm to limit the response of the corpus when transmitting messages, and the results show that the corpus scenarios have different network behaviors between them. This result ratifies the positive assessment of the corpus. From now on, the scientific community has a collection of scenarios where their routing algorithms and features can be tested, thus avoiding scenario selection, reducing time and eliminating unintended bias. The corpus contributes towards establishing a proper benchmarking scheme for OppNet routing algorithms where the routing performance is not relative to other routing algorithms but is examined overall.

5. Discussion

Nowadays, OppNet routing algorithms cannot be objectively evaluated nor compared because there is a lack of a globally accepted evaluation methods. This situation hinders the development of new routing algorithms. The present proposal intends to contribute toward an objective evaluation methodology by providing an analytically selected collection of scenarios, a corpus. This proposal will help ensure that evaluation results can be reliable, reproduced and contrasted in order to improve the objective quality.

Researchers have tried to evaluate their proposals fairly, for example, by evaluating each other's proposals, using scenarios that other researchers have used, or selecting metrics that fit their proposals. However, these evaluation approaches have not overcome problems such as lack of reproducibility or inability to generalize routing algorithms to any scenario.

It is common practice in OppNets to use well-known scenarios with a clear intention of standardizing evaluation methods. The problem, though, is not just a matter of using the same scenarios. If the routing algorithm being evaluated has to be general-purpose, it is also a requirement that the scenarios being used are representative of all possible network situations. Therefore, any collection of scenarios is not the solution, and what is needed is a fine selection of representative scenarios.

Besides the existence of a representative corpus, it is as well important that it be used by the community. The corpus introduced in this study has been proven as representative by means of experimentation, and has been made publicly available.

This work is not intended to create a dilemma of whether or not the corpus should replace the well-known scenarios. Obtaining a simple corpus is not a difficult task. There are different methods of obtaining a collection of scenarios in a straightforward manner, for example by using classical programming techniques such as random selection, trial

and error, genetic algorithms, or even machine learning approaches. However, obtaining a representative corpus is complex and challenging. A representative one represents, as a whole, all possible network behaviors. The selection of scenarios for a representative corpus goes beyond a cherry-picking process, and each scenario is carefully analyzed and compared with other scenarios. Still, the selection process may not matter as much as the corpus itself. The differences and representativeness of network behavior that the corpus has are what determines if a corpus is useful or not.

The corpus presented in this work was obtained via a creation methodology based on identifying the variables that characterize OppNet scenarios, methods to create OppNet scenarios and processes to assess differences and diversity among them. The differences and the representativeness of each scenario were carefully assessed. The results measured the representativeness and diversity of the corpus scenarios, showing significant differences. Therefore, it can be said that this is a representative corpus for objective evaluation. Having a representative corpus does not imply necessarily that it is the best. The scenarios of the corpus should be reviewed in the future, especially as new technologies emerge from arising new network behaviors.

The corpus comprises simple scenarios where network behaviors are uniform. There might be environments where it is interesting to have non-uniform behaviors, for example, when defining strategies where the routing algorithm changes depending on network conditions. These complex scenarios can be built, for instance, by concatenating simple scenarios from the corpus without unnecessarily expanding the number of scenarios in the corpus.

When there is a corpus, there is the risk of falling into the trap of developing tailored solutions that only work with the elements of this corpus. The behavior of a routing algorithm should not be finely adjusted to have an outstanding performance in each corpus scenario, since making a fine-tune would reduce the ability of a routing algorithm to extend the solution beyond the scenarios to the real world. Therefore, the routing model would not be able to generalize its routing abilities because the abilities would be too specific for the scenarios.

Another risk while developing routing algorithms for OppNets is to exclusively focus, or pay too much attention, to simulations using the corpus. Simulation is just a part of the developing methodology, which should always be followed by an emulation stage, testing with actual implementations of the algorithms, and real-world experimentation. Researchers should not overlook a complete methodology to convert a routing idea into real-world implementation.

6. Conclusions

From the state-of-the-art, in the review of the methodologies for creating routing algorithms, it was seen that, until now, there was no clear evidence to objectively evaluate and thus compare the performance of these algorithms. Evaluating and comparing routing algorithms is a complex task, and the final quality of the algorithm significantly relies on it.

To right this wrong, this study proposed a potentially global-agreed corpus for a fair evaluation and comparison of routing algorithms—a reference corpus of OppNet scenarios, which is a cornerstone in the design methodology. This corpus is a collection of forty-one methodologically obtained OppNet scenarios. These scenarios can be used to evaluate and thus compare the performance of routing algorithms. These scenarios were obtained using a creation procedure developed in this work that includes a backtracking process to enhance scenario diversity. This means that the corpus has the least number of scenarios, which, as a whole, represents most of the real-world OppNets.

Furthermore, for creating the corpus, it was necessary to characterize OppNets scenarios with a vector of characteristics. Such vectors are the basis for the analysis of similarities that lead to whether a scenario was a corpus member or not. The scenario is a node's contact trace described by a vector of seventeen characteristics. The corpus presented in

this work is a step toward creating a benchmarking scheme where the performance of routing algorithms is not relative to a selection of peers.

The corpus presented can be an important tool to help researchers follow the scientific method, especially regarding reproducibility and standardization aspects. These are essential features to improve quality research. The usefulness of the corpus requires that the community embraces it, using it for contrasting and evaluating routing performance results. The corpus is not static and should be revised to adapt to the needs; new technologies may require new scenarios in the future.

We look forward to this contribution simplifying and improving the development of routing algorithms in OppNets.

Author Contributions: Conceptualization, D.F., C.B. and S.R.; methodology, D.F. and S.R.; software, D.F. and C.B.; validation, D.F., C.B. and S.R.; formal analysis, D.F., C.B. and S.R.; data curation, D.F.; writing—original draft preparation, D.F.; writing—review and editing, D.F., C.B. and S.R.; visualization, D.F.; supervision, C.B. and S.R. All authors have read and agreed to the published version of the manuscript.

Funding: This work is partly funded by Secretaria de Educación Superior, Ciencia, Tecnología e Innovación (SENESCYT, ECUADOR), by the Catalan AGAUR 2017SGR-463 project, and by the Spanish Ministry of Science and Innovation TIN2017-87211-R project.

Data Availability Statement: The corpus is available to download at: <https://deic.uab.cat/~oppnet-corpus/> accessed on 30 August 2022.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Helgason, Ó.; Kouyoumdjieva, S.T.; Pajević, L.; Yavuz, E.A.; Karlsson, G. A middleware for opportunistic content distribution. *Comput. Netw.* **2016**, *107*, 178–193. [CrossRef]
2. Borrego, C.; Borrell, J.; Robles, S. Hey, influencer! Message delivery to social central nodes in social opportunistic networks. *Comput. Commun.* **2019**, *137*, 81–91. [CrossRef]
3. Chen, D.; Borrego, C.; Navarro-Arribas, G. A Privacy-Preserving Routing Protocol Using Mix Networks in Opportunistic Networks. *Electronics* **2020**, *9*, 2–15. [CrossRef]
4. Sarros, C.A.; Demiroglou, V.; Tsaoussidis, V. Intermittently-connected IoT devices: Experiments with an NDN-DTN architecture. In Proceedings of the 18th Annual Consumer Communications & Networking Conference (CCNC), Las Vegas, NE, USA, 9 January 2021.
5. Danielis, P.; Karlsson, G. Survey of mobile opportunistic networks for parallel data dissemination and processing. *KuVS-Fachgesp* **2020**, *1*, 1–3.
6. Nayyar, A.; Batth, R.S.; Ha, D.B.; Sussendran, G. Opportunistic networks: Present scenario—a mirror review. *Int. J. Commun. Netw. Inf. Secur.* **2018**, *10*, 223–241. [CrossRef]
7. Borrego, C.; Castillo, S.; Robles, S. Striving for sensing: Taming your mobile code to share a robot sensor network. *Inf. Sci.* **2014**, *277*, 338–357. [CrossRef]
8. Conti, M.; Giordano, S. Mobile ad hoc networking: Milestones, challenges, and new research directions. *IEEE Commun. Mag.* **2014**, *52*, 85–96. [CrossRef]
9. Trifunovic, S.; Kouyoumdjieva, S.T.; Distl, B.; Pajevic, L.; Karlsson, G.; Plattner, B. A decade of research in opportunistic networks: Challenges, relevance, and future directions. *IEEE Commun. Mag.* **2017**, *55*, 168–173. [CrossRef]
10. Mordacchini, M.; Passarella, A.; Conti, M. A social cognitive heuristic for adaptive data dissemination in mobile Opportunistic Networks. *Pervasive Mob. Comput.* **2017**, *42*, 371–392. [CrossRef]
11. Du, Z.; Wu, C.; Chen, X.; Wang, X.; Yoshinaga, T.; Ji, Y. A VDTN scheme with enhanced buffer management. *Wirel. Netw.* **2020**, *26*, 1537–1548. [CrossRef]
12. Borrego, C.; Amadeo, M.; Molinaro, A.; Mendes, P.; Sofia, R.C.; Magaia, N.; Borrell, J. Forwarding in opportunistic information-centric networks: an optimal stopping approach. *IEEE Commun. Mag.* **2020**, *58*, 56–61. [CrossRef]
13. Magaia, N.; Sheng, Z. ReFIoV: A novel reputation framework for information-centric vehicular applications. *IEEE Trans. Veh. Technol.* **2018**, *68*, 1810–1823. [CrossRef]
14. Tsaoussidis, V.; Borrego, C. Network Working Group P. Mendes, Ed. Internet-Draft Airbus Intended Status: Experimental R. Sofia Expires: 19 March 2021 fortiss GmbH 2020. Available online: <https://www.ietf.org/archive/id/draft-mendes-icnrg-dabber-05.pdf> (accessed on 25 April 2022).
15. Rajeswari, S.R.; Seenivasagam, V. Comparative study on various authentication protocols in wireless sensor networks. *Sci. World J.* **2016**, *2016*, 6854303. [CrossRef] [PubMed]

16. Kuppusamy, V.; Thanthrige, U.M.; Udugama, A.; Förster, A. Evaluating forwarding protocols in opportunistic networks: Trends, advances, challenges and best practices. *Future Internet* **2019**, *11*, 113. [CrossRef]
17. Sachdeva, R.; Dev, A. Routing in Opportunistic Networks: Implementation and Research Challenges. *J. Engg. Res. Icarl Spec. Issue* **2021**, *173*, 183. [CrossRef]
18. Alajeely, M.; Doss, R.; Ahmad, A. Routing Protocols in Opportunistic Networks—A Survey. *Iete Tech. Rev.* **2018**, *35*, 369–387. [CrossRef]
19. Vahdat, A.; Becker, D. *Epidemic Routing for Partially Connected Ad Hoc Networks*; Technical Report CS-200006; Duke University: Durham, NC, USA, 2000.
20. Lindgren, A.; Doria, A.; Schelen, O. Probabilistic routing in intermittently connected networks. In Proceedings of the International Workshop on Service Assurance with Partial and Intermittent Resources, Fortaleza, Brazil, 6 August 2004.
21. Spyropoulos, T.; Psounis, K.; Raghavendra, C.S. Spray and wait: an efficient routing scheme for intermittently connected mobile networks. In Proceedings of the SIGCOMM05: ACM SIGCOMM 2005 Conference, Philadelphia, PA, USA, 26 August 2005.
22. De Oliveira, E.C.; De Albuquerque, C.V. NECTAR: A DTN routing protocol based on neighborhood contact history. In Proceedings of the SAC09: The 2009 ACM Symposium on Applied Computing, Honolulu, HA, USA, 8 March 2009.
23. Grasic, S.; Lindgren, A. Revisiting a remote village scenario and its DTN routing objective. *Comput. Commun.* **2014**, *48*, 133–140. [CrossRef]
24. Kaner, C.; Bond, W.P. Software engineering metrics: What do they measure and how do we know. In Proceedings of the 10th International Software Metrics Symposium, Chicago, IL, USA, 11 September 2004.
25. Grasic, S.; Lindgren, A. An Analysis of Evaluation Practices for DTN Routing Protocols. In Proceedings of the Seventh ACM International Workshop on Challenged Networks, Istanbul, Turkey, 22 August 2012.
26. Petz, A.; Enderle, J.; Julien, C. A framework for evaluating dtn mobility models. In Proceedings of the 2nd International Conference on Simulation Tools and Techniques, Rome, Italy, 6 March 2009.
27. Sandulescu, G. Resource-Aware Routing in Delay and Disruption Tolerant Networks. Ph.D. Thesis, University of Luxembourg, Luxembourg, 2011.
28. Angius, F.; Gerla, M.; Pau, G. Bloogo: Bloom filter based gossip algorithm for wireless NDN. In Proceedings of the ACM Workshop on Emerging Name-Oriented Mobile Networking Design-Architecture, Algorithms, and Applications, Hilton Head, CA, USA, 11 June 2012.
29. Freire, D.; Robles, S.; Borrego, C. Towards a Methodology for the Development of Routing Algorithms in Opportunistic Networks. In Proceedings of the The Sixteenth International Conference on Wireless and Mobile Communications ICWMC 2020, Oporto, Portugal, 19 October 2020.
30. Zhang, Y.J. Evaluation and comparison of different segmentation algorithms. *Pattern Recognit. Lett.* **1997**, *18*, 963–974. [CrossRef]
31. Abdelkader, T.; Naik, K.; Nayak, A.; Goel, N.; Srivastava, V. A performance comparison of delay-tolerant network routing protocols. *IEEE Netw.* **2016**, *30*, 46–53. [CrossRef]
32. Bajaj, L.; Takai, M.; Ahuja, R.; Tang, K.; Bagrodia, R.; Gerla, M. *Glomosim: A Scalable Network Simulation Environment*; UCLA Computer Science Department Technical Report; UCLA Computer Science Department: Los Angeles, CA, USA, 1999; 990027, pp. 1–12.
33. Varga, A. OMNeT++. In *Modeling and Tools for Network Simulation*; Frederiksen, N.O., Gulliksen, H., Eds.; Springer: Berlin/Heidelberg, Germany, 2010; pp. 35–59.
34. Fall, K.; Ott, J. Delay-Tolerant Networking Research Group-DTNRG; 2002. Available online: <https://www.ietf.org/proceedings/75/DTNRG.html> (accessed on 20 May 2020)
35. Su, J.; Scott, J.; Hui, P.; Crowcroft, J.; Lara, E.D.; Diot, C.; Goel, A.; Lim, M.H.; Upton, E. Huggle: Seamless networking for mobile applications. In Proceedings of the International Conference on Ubiquitous Computing, Innsbruck, Austria, 16 September 2007.
36. Keränen, A.; Ott, J.; Kärkkäinen, T. The ONE Simulator for DTN Protocol Evaluation. In Proceedings of the 2nd International Conference on Simulation Tools and Techniques, Rome, Italy, 2 March 2009.
37. Riley, G.F.; Henderson, T.R. The ns-3 network simulator. In *Modeling and Tools for Network Simulation*; Frederiksen, N.O., Gulliksen, H., Eds.; Springer: Berlin, Germany, 2010; pp. 15–34.
38. Papanikos, N.; Akestoridis, D.G.; Papapetrou, E. CRAWDAD Toolset Tools/SIMULATE/uo/Adyton (v. 2016-04-21). Available online: <https://crawdad.org/tools/simulate/uo/adyton/20160421> (accessed on 23 March 2022).
39. Ciobanu, R.I.; Marin, R.C.; Dobre, C. Mobemu: a framework to support decentralized ad-hoc networking. In *Modeling and Simulation in HPC and Cloud Systems*; Joanna, K., Florin Pop, C.D., Ed.; Springer: Berlin/Heidelberg, Germany, 2018; pp. 87–119.
40. Kotz, D.; Henderson, T.; Abyzov, I.; Yeo, J. CRAWDAD Dataset Dartmouth/Campus (v. 2009-09-09). Available online: <https://crawdad.org/dartmouth/campus/20090909> (accessed on 1 August 2022).
41. Thiebaut, D.; Wolf, J.L.; Stone, H.S. Synthetic traces for trace-driven simulation of cache memories. *IEEE Trans. Comput.* **1992**, *41*, 388–410. [CrossRef]
42. Manfredi, V.; Crovella, M.; Kurose, J. Understanding stateful vs stateless communication strategies for ad hoc networks. In Proceedings of the 17th annual international conference on Mobile computing and networking, Las Vegas, NE, USA, 19 September 2011.

43. Souza, C.; Mota, E.; Manzoni, P.; Cano, J.C.; Calafate, C.T.; Hernández-Orallo, E.; Tapia, J.H. Friendly-drop: A social-based buffer management algorithm for opportunistic networks. In Proceedings of the 2018 Wireless Days (WD), Dubai, United Arab Emirates, 3 April 2018.
44. Borrego, C.; Borrell, J.; Robles, S. Efficient broadcast in opportunistic networks using optimal stopping theory. *Ad Hoc. Netw.* **2019**, *88*, 5–17. [CrossRef]
45. Cabrero, S.; Garcia, R.; García, X.G.; Melendi, D. CRAWDAD Dataset Oviedo/Asturies-er (v. 2016-08-08). Available online: <https://crawdad.org/oviedo/asturies-er/20160808> (accessed on 23 June 2022).
46. Bracciale, L.; Bonola, M.; Loreti, P.; Bianchi, G.; Amici, R.; Rabuffi, A. CRAWDAD Dataset roma/taxi (v. 2014-07-17). Available online: <https://crawdad.org/roma/taxi/20140717> (accessed on 3 March 2022).
47. Piorkowski, M.; Sarafijanovic-Djukic, N.; Grossglauser, M. CRAWDAD Dataset Epfl/Mobility (v. 2009-02-24). Available online: <https://crawdad.org/epfl/mobility/20090224> (accessed on 22 August 2022).
48. Akestoridis, D.G. CRAWDAD Dataset Uoi/Haggle (v. 2016-08-28): derived from cambridge/haggle (v. 2009-05-29). Available online <https://crawdad.org/uoi/haggle/20160828/one> (accessed on 23 June 2022).
49. Islam, M.R.; Rajon, S.A. On the design of an effective corpus for evaluation of Bengali Text Compression Schemes. In Proceedings of the 2008 11th International Conference on Computer and Information Technology, Khulna, Bangladesh, 27 December 2008.
50. Usama, M.; Malluhi, Q.M.; Zakaria, N.; Razzak, I.; Iqbal, W. An efficient secure data compression technique based on chaos and adaptive Huffman coding. *Peer -Peer Netw. Appl.* **2021**, *14*, 2651–2664. [CrossRef]
51. Arnold, R.; Bell, T. A corpus for the evaluation of lossless compression algorithms. In Proceedings of the DCC'97 Data Compression Conference, Snowbird, UT, USA, 25 March 1997.
52. Karamshuk, D.; Boldrini, C.; Conti, M.; Passarella, A. Human mobility models for opportunistic networks. *IEEE Commun. Mag.* **2011**, *49*, 157–165. [CrossRef]
53. Sandulescu, G.; Nadjm-Tehrani, S. Opportunistic DTN routing with window-aware adaptive replication. In Proceedings of the 4th Asian Conference on Internet Engineering, Pattaya, Thailand, 18–20 November 2008; pp. 103–112.
54. Yuan, P.; Wang, C. OPPO: An optimal copy allocation scheme in mobile opportunistic networks. *Peer -Peer Netw. Appl.* **2018**, *11*, 102–109. [CrossRef]
55. Schurgot, M.R.; Comaniciu, C.; Jaffres-Runser, K. Beyond traditional DTN routing: Social networks for opportunistic communication. *IEEE Commun. Mag.* **2012**, *50*, 155–162. [CrossRef]
56. Settawatcharawanit, T.; Yamada, S.; Haque, M.E.; Rojviboonchai, K. Message dropping policy in congested social delay tolerant networks. In Proceedings of the 2013 10th International Joint Conference on Computer Science and Software Engineering (JCSSE), Khon Kaen, Thailand, 29–31 May 2013; pp. 116–120.-
57. Bhattacharjee, S.; Roy, S.; Ghosh, S.; DasBit, S. Exploring the impact of connectivity on dissemination of post disaster situational data over DTN. In Proceedings of the 18th International Conference on Distributed Computing and Networking, Delhi, India, 3–5 July 2017; pp. 1–4.
58. Boldrini, C.; Conti, M.; Passarella, A. Social-based autonomic routing in opportunistic networks. In *Autonomic Communication*; Springer: Berlin/Heidelberg, Germany, 2009; pp. 31–67.
59. Freedman, D.; Pisani, R.; Purves, R. *Statistics (International Student Edition)*, 4th ed.; WW Norton & Company: New York, NY, USA, 2007.