An Analysis of Evaluation Practices for DTN Routing Protocols

Samo Grasic
Luleå University of Technology
{samo.grasic@ltu.se}

Anders Lindgren
Swedish Institute of Computer Science
{andersl@sics.se}

ABSTRACT
Over the past decade, a lot of research has been done to develop efficient routing protocols for Delay- and Disruption Tolerant Networks (DTNs). In the course of this work, many comparative evaluation studies have been done to determine which of two proposed protocols is the better one (for a given situation). The majority of these evaluations are based on results gained from simulated network environments. In order to conduct a relevant evaluation of routing schemes, numerous conditions, policies and data need to be specified and fed into the simulation environment. The aim of our work in this paper is to discuss current DTN evaluation practices through a thorough and critical literature study. Based on the surveyed material, we show some weaknesses and lack of argumentation used in the evaluations. Through this, we hope to aid in bridging the gap between simulated and real-world DTN environments. In addition, and as a call for further research, we propose a model for evaluation of DTN routing schemes that outlines the most crucial inputs that needs to be considered in the evaluation process. This model is then projected onto evaluation practices used for evaluations of DTN routing protocols in a set of sixteen papers roughly covering a decade of DTN research.

Categories and Subject Descriptors
C.2.2 Network Protocols [Routing protocols]:

General Terms
Algorithms, Design, Experimentation

Keywords
DTN, routing, evaluation practices

1. INTRODUCTION
Routing of network traffic is one of the main challenges in the field of DTN research. The use of efficient routing schemes is a key factor for good performance of DTN networks. Therefore, the usability and applicability of DTN deployments are conditioned by efficient and optimized routing algorithms. Over the last decade, numerous routing protocols have been proposed within the DTN research community. Although the DTN architecture originates from interplanetary internet space research, it is widely applicable for terrestrial use case scenarios [1] of challenged networks. Real-world deployments of DTN in hard-to-access remote areas or creation of large scale test-beds are however costly and hard to control. Therefore, a majority of DTN research has been conducted in laboratories and simulated environments.

Over the past decade of DTN research, researchers have gradually increased the complexity of the simulated environments in order to make them more realistic. Development of complex synthetic mobility models [2] and increased collection of mobility and connectivity traces from real-world networks [3, 4] have allowed researchers to perform better DTN routing scheme evaluations.

However, as important as the choices of routing protocol, mobility pattern, and network connectivity are to the performance of the system, there are also other evaluation factors that needs to be considered. Such inputs as routing protocol parameters, appropriate time scales for a given scenario, network traffic, and the validity of the simulator code are often not discussed at any length when routing protocols are evaluated. This is troubling as they can have a significant effect on the performance of the evaluated routing schemes.

In this paper, we first present an overview work containing evaluations of DTN routing protocols in the literature through a survey conducted over a large body of research papers. We then propose a simple DTN evaluation model that can be used a guideline in the evaluation process for when setting up DTN environments for tests. Afterwards, the proposed evaluation model is used to study the work in the literature from some different aspects regarding evaluation methodology and validity. This finally lets us draw some conclusions regarding current evaluation practices and point out some potential weaknesses.

2. LITERATURE SURVEY
We have studied a large body of work from current and past research in DTN routing protocols and evaluation studies of such protocols. For this study, sixteen papers published in scientific journals and conferences were surveyed. The selected set of papers present new or improved DTN routing schemes and/or perform evaluations of some proposed protocols. The selected papers were chosen to include the most well-known and cited DTN routing protocols and to cover publications spanning about a decade of DTN research with on average two papers per year. This set contains the following papers: [5, 6, 7, 8, 9, 10, 3, 11, 12, 13, 14, 15, 16, 17, 18, 19].

We are well aware that the selected set papers does not cover all research that has been done regarding DTN routing over the past
decade. This would not be possible due to the vast amount of papers in this area. However, we believe that this set of papers provide a representative selection of the work that has been done in this area, and can be used to critically reflect the scientific practices used in the evaluation of DTN routing protocols. The intention of this paper is to offer a qualitative study of evaluation practices currently in use, therefore the number of papers in the analysis is intentionally kept low.

The authors would like to stress that the aim of this work is not to question the theoretical or evaluation validity of any individual papers used in this set. The intention is rather to create awareness and spark discussions regarding the methods used in this research area.

The rest of this section will provide a brief chronological overview of the papers included in the survey, and subsequent sections will address concerns in the evaluation practices from different aspects.

2.1 Chronological Overview

One of the first papers in this research area was published in 2000, when Vahdat and Becker proposed the Epidemic Routing protocol [5] for partially connected ad hoc networks. The paper discusses three concrete motivating real-world scenarios such as mobile sensor networks, natural disaster recovery situations and military deployments, where the protocol is applicable.

Eventually, more research happened in this and related areas and the term Delay Tolerant Networking (DTN) was coined [1]. As this happened, the PRePHET routing protocol [6] was proposed. Similarly to the Epidemic Routing paper, the authors dedicate a portion of the introduction of the paper to a discussion of motivating scenarios (in this case, the protocol applicability for the nomadic Saami population in Arctic part of Sweden and rural villages in India was discussed). In 2004, Jain et al. proposes a couple of deterministic routing schemes [7]. An example of connecting the remote village with the DTN was used and discussed. Leguay in 2005 describes routing scheme for a mobility pattern space [8] using reference to a mobility case scenario of students in a campus and corporate environment. In the same year spray and wait routing scheme [9] was proposed. Applicability of the protocol was argued with a few references to a wildlife tracking, military networks and nomadic networks. MaxProp routing scheme [10] introduced in 2006 was designed and evaluated on their own vehicle-based UMassDieselNet DTN test bed. RAPID routing scheme [3] was studied on same test bed, but extended test bed in the same year. Social based routing algorithm BUBBLE Rap [11] was proposed in 2008. The algorithm was applied to the gathered experimental datasets of human mobility in the cities and academic environment. The ORWAR routing scheme [12] proposed introduced in the same year uses references to a couple of scenarios, but it does not explicitly discuss its applicability. In a similar way, the Encounter-Based routing scheme [13] was presented, giving only references to numerous papers discussing challenging environments. In 2009, de Oliveira et al. proposed the NECTAR routing scheme [14] that exploits the neighborhood contact history and was applied to the mobility of the students in the university campus. E-PRePHET routing scheme [15] introduced in the same year by only mentioning applicability of the protocol with several references. Use of DTN in urban transport system was well discussed in 2010 together with the RUTS routing scheme proposal [16]. In the paper proposing an adaptive scheme switching between the AODV routing protocol and DTN routing [17] from the same year, the design was discussed only on the abstract level of MANETs and DTNs. However the San Francisco cab and Helsinki synthetic traces were used in the evaluation. One year later in 2011 the R3 routing scheme [18] was introduced, targeting mixed DTN and MANET scenarios using the references for extending the reach of Internet, wildlife monitoring, content distribution and tactical operations. The last paper from the selected set of papers proposes the Weak State Routing protocol [19] in 2011, referring to the applicability of it in the rural areas and in urban vehicle scenarios.

3. DTN Evaluation Model

In order to evaluate a DTN routing scheme in a simulator, it is necessary to implement the routing algorithm and run it in the simulated DTN environment. While running the simulation, different kinds of network performance metrics are collected and saved for further analysis and interpretation of the results. In this section, we consider the different inputs and outputs that are relevant for evaluating a DTN routing protocol and propose a simple conceptual model.

According to Jain et al. [7], it is not immediately obvious what the most desirable objective for a DTN routing protocol is and it cannot be defined by one simple metric. In traditional fixed networks, the objective of an intra-domain routing protocol is usually to minimize the number of hops traversed by a packet. In addition to minimizing the hop-count (which also gives an indication of the overhead of the routing protocol), the objectives of a DTN routing protocol is also likely to include maximizing the message delivery probability and minimizing end-to-end latency. Even though these metrics are closely related to each other, we will use them all as output from our evaluation model since they are helpful in interpreting the routing scheme performance.

Besides the routing algorithm itself, the outcome of a DTN routing protocol evaluation also depends on the scenario in which the evaluation is done. Spyropoulos et al. [20] outlines network connectivity, mobility of nodes, node resources, and application requirements as the set of main characteristics relevant to routing. They are directly affecting the routing performance and should all be considered in the routing scheme evaluation process. Keränen et al. [21] investigated the characteristics of synthetic mobility models and showed how mobility and connectivity models affected the performance of different routing schemes. Karvo and Ott [22] investigated the importance of using of the proper time scales and the correct protocol parameters in DTN routing protocol evaluations. They show that these settings need to be chosen so that they correspond to the mobility characteristic of the network nodes. In their paper, they show that different results can be obtained by using the same evaluation scenario and changing only the time scales and values of the used protocol parameters. The DTN performance is also conditioned on the definition of hardware resources of network nodes, such as storage capacity, radio link range and bandwidth, As Balasubramanian et al. show in their work [3] the DTN routing objective can be defined as a resource allocation problem.

As outlined above, it is crucial that the DTN routing scheme evaluation results are interpreted vis-à-vis the used evaluation scenario. Therefore our DTN evaluation model considers Routing Protocol, Routing protocol parameters, Connectivity and Mobility, Network Traffic, Node resources and Time scale as important inputs into the evaluation process. Figure 1 shows a simple DTN evaluation model that can be used as a guideline for future DTN routing scheme evaluations. Our aim with this model is not to make all evaluation scenarios identical, but to motivate DTN researchers to put more effort into discussing their selection of inputs into the simulated environment. The inputs to the model are of course not completely independent as there is, for example, a strong correlation between mobility and connectivity patterns, and the appropriate time scales to use is often dependent on mobility and other user behavior. There-
fore, researchers should take these correlations into consideration and also discuss the impact the different inputs have on each other. As researchers have different application scenarios in mind when they are developing routing schemes, different DTN network scenarios will be used in their evaluations. This high variance of evaluation conditions and scenarios may make a straightforward comparison of different evaluation studies rather difficult. Therefore, even stronger argumentation about the selected evaluation scenario is needed in order to understand and interpret the results. By doing this it is possible to make the DTN evaluations more comparable and provide clearer interpretation of the results for the reader.

4. MOBILITY AND CONNECTIVITY

4.1 Real World Applicability
A real-world contextualization of DTN research can provide motivation and validate the chosen mobility and connectivity models used in the evaluation process. As can be seen from the literature, researchers have different kinds of real world DTN applications in mind to motivate their work.

Chronological reading of the selected papers offers an insight into how the amount of text and arguments motivating the research for a particular routing scheme fades out throughout the years. Papers from early DTN research [5, 6, 7] contain wider discussions of the applicability including examples of concrete real-world DTN scenarios. Papers published later are usually motivating their research only with citations of previous work. As much as this is valid scientific practice, authors often do not return to these references [8, 15, 17] in their discussions and leave the reader in the abstract world. Exceptions are papers based on novel concepts in networking scenarios not used before, such as social networking [11], vehicle networks [7, 10, 3] or combination of DTN and MANET scenarios [18, 19] in our case.

4.2 Network Density
Throughout all the papers the following real-world DTN applications are motivated as potential scenarios: connectivity with the remote rural village, wildlife sensors networks, natural disaster communication scenarios and military applications. A common characteristic that can be assumed for all these scenarios is that they cover large areas and the number of nodes in the network can be relatively small. Therefore, highly partitioned and sparse networks can be expected in these kind of scenarios. In contrast to this, the non-abstract evaluation scenarios of the examined body of work are usually placed in densely populated urban areas such as cities [6, 7, 10, 3, 12, 13, 16], conference centers or university campus areas [11, 14, 18, 19]. Use of the dense network scenarios with the high number of nodes (up to 2000 [14]) forces researchers to analyze evaluation results using only quantitative and statistical methods. By following some real-world DTN deployments it can be seen that the number of nodes in the DTN hardly reached 40. In the N4C project deployment in rural areas [23], after 3 years of DTN deployments in the rural villages the number of used nodes hardly reached 20. Around 40 buses were used in the UMass Diesel Net deployment [3]. This points out that there is a need for investigating DTN routing schemes behaviors on the downscaled scenarios.

In the discussion section of the AODV routing scheme [17] authors put forward an observation that the AODV did not perform as good using when using the real-world traces that it did with the Random Waypoint scenarios. However significant effect this issue can have in the real-world deployment, discussion is concluded with the expectation that the some of the real-world scenarios will be more clustered.

As much as it is easy to understand that every researcher hopes that the proposed routing scheme will become widely used in a scaled up real-world scenario, using only large scale scenarios can also hinder the growth and use of the DTN from the smaller scales.

4.3 Hop-count
Hop-count is a metrics in DTN evaluations that refers to the number of nodes through which message has to pass between source and destination node. This metric is important for the interpretation and analysis of the routing performance. Together with the message delivery rate and average delivery delay, it helps to understand how long routes messages had to pass from source to destination or how much network resources were used, etc. Therefore the information about the average hop-count can speak about the how
challenging the evaluating scenarios was and how well the routing scheme managed to utilize network resources.

For example, Sandulescu et al. in their work [12] do not provide results about the average number of hops required to reach the destination in their evaluation. However, through close examination of the provided evaluation results of the Direct Delivery routing scheme [21], it can be seen that almost 60% of the generated network traffic could be directly delivered from the source node to the destination node. In other words, only 40% of the generated network traffic in their scenario required network routing in order to be delivered.

Although the information about needed number of hops in the routing evaluations speaks about the routing challenge, only half of the examined body of work provide evaluation results that include information about the average number of hops needed [5, 6, 10, 11, 14, 15, 17, 18].

4.4 Proper Mobility Scenario

DTN routing schemes are using different kinds of methods and algorithms to route the network traffic over the network. In order to reach optimal network performance, a majority of routing schemes tries to leverage the underlying characteristics of network mobility and connectivity. Therefore, it is important to carefully choose the mobility and connectivity data that matches the routing scheme in the DTN evaluations.

The simulated environment is fed with connectivity or mobility data that can be collected from the real-world deployment or synthetically generated data from the model. Use of synthetically generated mobility and connectivity data is popular within the DTN community. In our examined empirical material more than three quarters of evaluations used synthetic mobility models.

The E-ProPHET routing protocol [15], similarly to ProPHET [6], uses the history of previous node encounters, contact frequency and contact duration. Despite the fact that both protocols are relying on predictable and repeating node mobility, the protocols were evaluated and compared with Random Way-Point (RWP) mobility model [21] that does not contain repetitive mobility of the nodes that the protocol could leverage.

This is common practice when several different DTN routing schemes are evaluated and compared. For instance, history based routing schemes such as MaxProp [10] or ProPHET [6] are often compared with other routing schemes in evaluations based on the RWP mobility model [3, 13, 15]. Using inappropriate mobility models in DTN routing protocol evaluations can significantly affect the routing protocol performance. Although authors of the routing protocols usually chose the right and most favorable mobility model for their own routing protocol, it is somewhat unfair to evaluate routing protocols that were not designed for the selected mobility scenario.

Our critique here is not a claim that DTN routing protocol researchers are deliberately choosing the evaluation scenarios that favors their own proposed routing scheme. It is rather an observation of the lack of arguments and discussions regarding how and why a certain evaluation scenario and the set of protocols used for comparison in the evaluation was chosen. There are instances where a protocol has been claimed to perform poorly where the suboptimal performance has clearly been due to the fact that the evaluation scenario is very different from the type of network for which the protocol was designed. We do however still believe that it is useful to include routing protocols that are not designed for a particular scenario in evaluations. This can be beneficial and informative and provide a deeper understanding of the fundamental workings of the protocol that the authors did not consider. Such evaluations should on the other hand always include an explicit remark about this discrepancy between intended and used scenario so that the results for that protocol are not considered a general benchmark for all types of scenarios. It is worth to mention that in the examined body of work, only one paper [10] explicitly states and reminds that the presented results from evaluation are applicable only to the scenario used in that paper.

5. NETWORK TRAFFIC

The majorities of real-world DTN deployments were deployed and used as test bed and did not include actual users that would use the DTN for running their applications. Hence, in the contrast with wide set of mobility and connectivity traces available, real-world DTN data traffic traces are scarce. To us, the only known the real-world real-use DTN traffic traces that are available were collected in the N4C project deployment in Arctic village and involved more than 30 users [24, 25].

The scarcity of real-world DTN traffic traces with real users force researchers to use synthetic traffic generators. In all examined papers, synthetic traffic generators were used in the DTN routing scheme evaluations. The network traffic load is defined by the variance of message sizes and the rate of creation of messages per time unit. In many DTN routing protocol evaluations, the routing schemes are analyzed with the use of different network traffic loads.

All the evaluations use randomly generated destinations for generated network traffic. This Peer-To-Peer based communication assumption in the DTN routing scheme evaluations does not allow analysis of the network traffic congestion. There is also lack of sparsely connected network scenarios used in the evaluation where formations of network bottlenecks topologies are more expected.

An analysis of collected real-world DTN traces with real users in [23] showed that only 22% of the network traffic was sent among peers and almost 78% of the network traffic was sent from or to the gateway in the deployed DTN. Lindgren et al. [26] studied the performance bounds that can be achieved by opportunistic multi-hop forwarding in mobile networks based on real mobility traces. In one of these studies, a comparison was also made to see the difference between using synthetically generated traffic, and using traffic patterns that were based on the real communication patterns of the users in the trace. This had a major impact on the end result. Therefore, the effects of different kind of network traffic flows on the DTN routing scheme performances calls for future research as synthetic traffic models may not yield valid results.

Recently, there has been much work on trying to define realistic mobility models, but very little work has been done on determining what realistic communication patterns will look like in networks like these, which is troubling. To a large extent this is likely due to a scarcity of traces from systems with real users, and the traces that do exist, while useful, are usually limited to a small number of users or very specific scenarios. There is a need for further work in this area, and as long as there is a lack of common understanding on communication patterns, researchers need to take care to explain their choice of traffic models, what types of scenario it is likely to occur in, and which characteristics of the routing protocol it is likely to highlight and affect.

6. NODE CHARACTERISTICS

6.1 Storage buffer and bandwidth

The constrained DTN node characteristics are part of the routing objective definition. The foremost considered characteristic of network nodes are node’s message storage capacity and network
interface transfer speed. They are defined with the use of realistic storage units such as bytes or in more abstract terms such as the number of messages that the node can store. These node characteristics can only be read and understood with the rate and size of generated network traffic.

Although, varying node storage constrains and network traffic load is a popular method to analyze the routing performance, a high variance of node characteristic used in the DTN evaluations makes a comparison between evaluations almost impossible. The average message size varies from 100B to 1MB, network interface transfer speed varies from 3KB/s [7] up to 2MB/s [15] sometime even to infinite bandwidth [8]. In a similar vein the node storage capacity varies from 10KB to 40GB [3].

Interpretation of the results is even more problematic when some used network node characteristic are not mentioned in the evaluating text, e.g. not mentioning storage capacity [8] or not mentioning any of the node characteristics [16].

Setting up the right node characteristic that matches the evaluation scenario and routing objective is a challenging problem. Fast technological development of hardware components of the node is constantly changing the routing objective. Use of hardware characteristics that available during the research time can make the research results more applicable in real-world. For instance, setting up nodes’ storage capacity to 300KB [19] when the average low cost off the shelf router has tens or hundreds of megabytes of available memory does require some additional reasoning.

For instance authors of the R3 routing scheme mention in their work [18] the assumption and expectation of loosely synchronized clocks across nodes in the network. Although time synchronization within the DTN is known to be a challenge [1, 27, 28] and the synchronized time is crucial for the R3 protocol performance, the time synchronization tolerance is not defined or discussed in their work. However this might be a serious real world deployment issue, such weaknesses of the routing schemes are usually not explicitly mentioned or discussed.

7. PROTOCOL PARAMETERS

Almost every DTN routing scheme has its own specific parameters that allows adopting the routing algorithm to the specific used case scenario. With the use of the default or wrong set of routing scheme parameters the routing performance can be seriously hindered. For instance if the time scale used in the evaluation scenario does not match the parameter that defines the aging speed of the history based routing algorithm, all the routing decisions will be based on the invalid knowledge. Hence, the used protocol will not be able to perform at all.

It is a common practice that scientist discuss only how they have set up parameters for their own routing protocol. Usually routing parameters of other routing protocols used in the evaluations are only mentioned and set to the default value that does not necessary match the scenario. In some cases [16], the parameters of other protocols used in the evaluations are not even mentioned, leaving the reader with only the hope that at least the default values of parameters were used.

7.1 Time Scales

The importance of the DTN routing evaluation time scale was already mentioned in the discussion about routing scheme parameters related to time unit such as aging rate of data in the history based routing schemes [6, 10, 15].

In scenarios when adaptive and history based routing schemes are used it is important to consider needed warm-up time, that allows routing schemes to adapt to the network scenarios. How and if the warm-up time was considered in the evaluation usually remains an open question, since it is rarely explicitly mentioned in the examined evaluations.

Simulation times of DTN evaluations in the examined body of work varies from less than an hour [15] up to eight months [11]. A very closely bound parameter that defines the routing objective itself is the expiry time of sent messages. In a similar vein as the simulation time, set message expiry time has a high variance in the evaluations from 20 seconds [3] up to three weeks [11]. Although the information about message expiry time is crucial for the interpretation of the results it is not always stated [15].

8. CODE VALIDITY AND AVAILABILITY

The majority of the routing scheme evaluations are conducted in the simulated environments. In more than one third of the examined papers [6, 7, 8, 9, 10, 3] researchers developed and implemented their custom simulator. In cases when external mobility traces were not used they also had to implement their own mobility model. In another third of the work researchers used popular NS2 simulator [5, 14, 17, 19] widely used in the wireless networking research, the Huggle simulator [11] or QualNet simulator [18]. The rest of work [12, 13, 15, 16] are based on the ONE simulator [21] that was designed for evaluation of DTN routing scheme and it has become popular within DTN research community. With only one exception [7], authors that used their custom simulator implementation do not mention or discuss the validity or availability of used code.

In a similar vein the availability of the routing algorithm code implementation is not mentioned or discussed in any of the papers except [7]. This issue is even more problematic when the researcher are implementing other routing schemes, whose implementation details are not that familiar as they are for their own routing schemes.

How problematic this issue can be, have authors of this paper discovered when they wanted to investigate after a poor PROPHET performance reported in [3]. After lengthy process of retrieving the custom simulator source code from authors (with their full cooperation), close examination of the PROPHET Source code was conducted. Despite the vast effort put in the process of implementing the simulator and other routing protocols, two bugs were founded in the code that completely hindered the PROPHET routing scheme. Although the sharing the source code with the research community does not bring the code validity per se, it shorten the code retrieval time and can help to increase the code validity in a long run. With the presence of The ONE simulator and researchers contributing their routing scheme implementations a great step forward was done that allows researchers to use the other routing schemes for comparison and elaborate more on the synthetic mobility models. However, more recent trend noticed in more recent works [18, 19] in the DTN research points out towards the hybrid scenarios of MANETs and DTNs for which The ONE simulator was not designed. Hence, new tools and simulators might need to be developed.

9. DISCUSSION AND CONCLUSIONS

As shown in previous sections, DTN routing protocol evaluation results that only provide some commonly used metrics, such as delivery ratio, average delay or overhead ratio, are not sufficient. Such “results” in themselves without further context and interpretation cannot provide a good understanding of the more fundamental workings of the routing protocols and the reason behind those particular results.
Evaluation of DTN routing is a multidimensional problem that requires consideration of all types of inputs into the simulator. The simulated DTN environment is giving DTN researchers a chance to perform controlled and easily reproducible experiments. At the same time, it challenges the researchers' understanding of the system and target scenario they are trying to evaluate the protocol within. In order to set up an evaluation scenario in a simulator that is applicable to real-world DTN deployments, the fundamentals of the targeted deployment scenario must be understood. As a result of the wide spectrum of potential uses for DTN protocols, a high diversity of evaluation scenarios can be found in the DTN routing research literature. However, in order to benefit from this diversity, it is crucial to put forth strong arguments for the scenario design choices made, and to provide exact definitions of used simulated scenario in research papers so that the reader can reproduce the results and also intuitively get an understanding for the types of scenarios where the presented results are applicable.

A similar problem in the existing work is that there is often a strong correlation between different aspects of an evaluation scenario as can be seen in Figure 2. Scenarios span a wide spectrum of realism from completely synthetic scenarios (like random waypoint) to very realistic (often based on trace data), but this is then also commonly correlated with how easy applicable the results are in other types of scenarios. The evaluations performed in very realistic settings are often very specific to that particular scenario and any conclusions drawn cannot be expected to be valid in other scenarios. On the other hand, more uniform and synthetic scenarios might lack some realism but can instead highlight some more general aspects of the evaluated protocols. The ideal target scenario would be one that realistic and at the same time provides generally valid results. Such scenarios are however unlikely to attain, so instead researchers should take care to specify where a scenario is located within this space.

The evaluation model proposed in Section 3 could be used as a starting point to provide guidelines on how to conduct DTN routing protocol evaluations in future DTN research. Even though the inner workings of the simulated code can often be a “black box”\(^1\), the inputs and outputs given in the model should be attempted to be specified whenever reporting on the results. We claim that close consideration of outlined evaluation process inputs can bring more consistent and readable evaluations into the DTN research field. With the awareness of the beneficial diversity of evaluation scenarios we are not advocating generic DTN evaluations, but seeking for better argumentation when parameters and inputs are strongly deviating from the expected condition in concurrent real-world scenario. New and upcoming DTN routing challenges are already now challenging the proposed evaluation model. However, the evaluation model can be changed, developed and adopted throughout further argumentation and discussion.

Thorough reading of the empirical DTN evaluation material revealed that when proposing a new routing protocol, authors rarely dare to mention weaknesses in their routing scheme. The only space for self criticism could be found at the end of the papers, where authors are discussing future work plans. A separate chapter discussing routing scheme weaknesses and potential real-world implementation challenges should be encouraged in future DTN research, since the creators of a new routing scheme are probably the ones the most familiar with potential problems. Such input to the community could thus be helpful in guiding future research or deployment work based on that protocol.

10. REFERENCES


