Alternative Topology Construction for Cooperative Data Distribution in Mobile Ad Hoc Networks

Milan Cvjetkovic, Veselin Rakocevic
Department of Electrical and Electronic Engineering
School of Mathematics, Computer Science and Engineering
City University London, UK
Emails: {milan.cvjetkovic.1; veselin.rakocevic.1}@city.ac.uk

***Abstract*—Ad hoc networks of mobile nodes can be used to extend or improve connectivity, and cooperative data distribution represents the basis of this approach. Cooperation in such networks can be improved by providing nodes with information on network dynamics and topology changes. This paper proposes a proactive approach to handling mobility-induced network topology changes. The approach uses signal strength trends to predict the future locations and connectivity between the network nodes. Our research is aimed towards the creation of an alternative topology in each node, where a node would keep other nodes suitability for cooperation in the data distribution. In this paper we present the initial algorithm and test the prediction method using a simulation. The algorithm is then applied in an experimental testbed where its performance was tested using real moving nodes executing a real data distribution process. The performance results show a significant improvement in terms of file transfer delay.**

I. Introduction

With the growth of the number of services being shifted online and in real-time, faster, more reliable and further reaching networks are in higher demand. The way these services are accessed – anytime and from any location – is also increasing the need for wireless and mobile networks capable of performing as close as possible to the fixed networks. This also means that data should be distributed across the network in multiple copies and as close as possible to end users, as noted in [1].

Mobile ad hoc networks can be viewed as not only an affordable, instant and adaptable solution for connecting collocated mobile nodes, but as a mean for extending the fixed wireless networks [2], offering a solution for the above-mentioned increasing need for data availability. Existing MANET data distribution solutions [2], [3], [4], [5] rely on nodes cooperatively caching data and distributing it among themselves. Coordinating nodes' behaviour in a highly dynamic and completely decentralised and distributed environment, without an external controller or a reference point is one of the greatest challenges presented upon MANET cooperative data distribution.

Using a *reactive* model, where any change in the network topology induced by mobility can result in a broken link and an unsuccessful transmission, and only then recovering by discovering other sources or routes, consequently reduces the data distribution efficiency. If nodes were aware of the network dynamics, taking a *proactive* approach by predicting the future locations of other nodes' in relation to them, this would help maintaining a higher data distribution efficiency, by discovering alternative solutions before an error occurs or efficiency drops significantly.

Our research is aimed at developing an algorithm which would enable nodes to predict the changes in topology and evaluate potential cooperation partners, and make data distribution decisions based on this knowledge. This algorithm should take into account neighbouring nodes' parameters such as signal strength, trend of signal strength over time, mobility parameters such as speed and direction of movement, data distribution oriented related parameters such as cache content, neighbourhood stability and access to an external data source. As a result of measuring, collecting and analysing these parameters, every node calculates an evaluation of each neighbouring node's *suitability for cooperation in data distribution*.

The measured parameters for all the nodes in the network make up what we call an *alternative topology* of the network, representing not the physical or logical topology, but a measure of capacity for cooperation for each of the other nodes. This kind of a database has a goal of enabling nodes to cooperatively distribute data in a more efficient way.

This paper will focus on our approach to using signal strength trends to predict the future relation of the nodes, which is the first step towards building the full alternative topology. We will also apply this approach in an ad hoc network using an experimental testbed, to test its impact on data distribution efficiency.

The rest of this paper is organised as follows: the second section covers existing approaches to cooperative data distribution and mobility and signal strength prediction in decentralised networks. Our signal strength prediction approach and its simulation verification are described in the third section. The application of signal strength prediction to a cooperative data distribution ad hoc network and experimental testbed results are represented in the fourth section. The conclusion is given in the fifth section, alongside the future work.

II. Related work

Mobility issue in cooperative data distribution for MANETs

Mobile ad hoc networks face different types of challenges [2], some inherent from their wireless and infrastructureless nature, and some due to node mobility. The existing cooperative data distribution solutions for MANETs focus on mitigating the adverse effects caused by these challenges, by distributing the data throughout the network as part of each node's cache, making it more available and "closer" to each node.

For example, the work presented in [2] describes a co-operative caching scheme aiming to reduce cache miss ratio and hop-distance and the number of generated messages per successful request, while increasing the successful request ratio. This scheme uses a mix of flooding and hop-by-hop requests, organising nodes into cooperation zones. The authors of [3] propose a cooperative and adaptive caching system based on different types of nodes – caching nodes and query directories. The former are the nodes actually caching data received from external sources and the latter serve as registers of cached data locations, handling data requests. The aim of this mechanism is reducing delay and increasing hit ratio. Mechanism presented in [4] aims to increase throughput, reduce hop-distance needed for a request resolution and increase cache hit ratio, using broadcast flood requests and a caching policy based on proximity to a known source and time since last update. Cooperative caching solutions proposed in [5] have increasing cache hit ratio and reducing delay as their goal, and are relying on hop-by-hop directed data requests, while caching data, paths to data, or both, based on content popularity, hop-distance and TTL parameter.

These schemes differ among themselves in several parameters, such as cache management, resolving data requests, data popularity models, network models and the metrics used to evaluate. The one thing they do have in common is the fact that they are not aware of the changes in node locations and their mutual relations, until these changes result in an error in communication, i.e. they are taking a *reactive* approach to the dynamics of a MANET environment.

Link quality prediction

In an ideal scenario, nodes in a wireless network would have a complete knowledge of each others location and movement, enabling them to adapt and plan ahead, optimising the efficiency of the data transfer. To some extent, such a scenario is possible in infrastructure based, centralised network, where an omniscient central controller can direct traffic according to the location, movement and link quality information it has at its disposal. An example application of this scenario is inter-cellular handover in cellular mobile radio networks.

Achieving such, or a comparable, level of global knowledge about node location and mobility without using any external localisation references (such as GPS) in a decentralised ad hoc system, where nodes communicate only directly among themselves, and rely on others to propagate their messages, would be impossible or impractical. Even if that was not the case, determining the exact locations of a pair of nodes does not guarantee us knowledge about the link quality between them. On the other hand, nodes are capable of taking signal strength measures from other nodes on their own. This parameter is more valuable than determining actual location of the other nodes, as it more directly represents the potential for successful communication, as noted in [6].

This paper proposes a link quality prediction technique for wireless networks. The signal-to-noise levels are measured over time, and time patterns are compared by cross-correlating them to "training data". The set of past sequences whose cross-correlation with the measured sequence is above a set threshold is used as a base for prediction. Intervals following these past sequences are considered and compared among themselves, in order to find the one which is most likely to follow (the one with the highest average cross-correlation within this set) and use it as the prediction.

The rationale behind this technique is the fact that human movement follows certain patterns, especially in indoor environments. Also, the fact that similar mobility patterns produce similar link quality measures is exploited. This means that if there is a regularity in movement which can be predicted, the resulting SNR measures are also predictable.

III. Signal Strength Prediction

As previously stated, the alternative topology reflects other nodes' suitability for cooperation in data distribution, and signal strength, being related to data transfer speed, is an integral part of our alternative topology creation algorithm. This section describes our approach to signal strength prediction.

Our signal strength prediction mechanism is a very simple, intuitive one: if there is a trend in signal strength measures through time, a prediction is made that the signal strength will continue following that trend in the next measuring interval. This corresponds to detecting nodes getting closer or moving away from each other. We assume that in case of mobile nodes following human mobility patterns, the trends are stable enough to make the previous statement hold. This assumption was tested using a simulation, results and setup of which will be discussed later.

Signal strength measurements (ss) are taken by each node in regular intervals (Δt) for each other node and the monotonicity—whether the signal strength is increasing or decreasing—is followed, forming a *signal strength trend vector* **SST**, where:

$$\mathbf{SST}(t) = ss(t) - ss(t - \Delta t) \tag{1}$$

Prediction is done based on a simple linear model, where m represents memory, or the number of previous signal strength trend values taken into account. If signal strength has had a clear trend (increasing, stagnating or decreasing throughout the whole range of memory) in the past m steps, a prediction is made that the trend will continue, i.e. an increasing signal strength will continue to increase, a stagnating one will continue to stagnate and a decreasing one will continue to decrease.

Following expression represents the workings of our prediction model, where $p(t)$ represents the predicted value of signal strength change for time point t. :

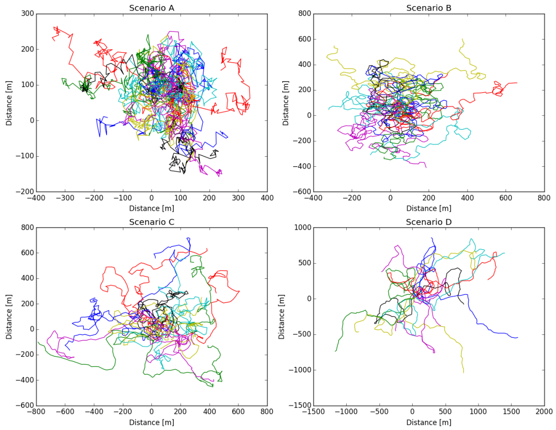

Fig. 1. Node paths for different RWP models

$$p(t) = \begin{cases} 1^{1} \;\; if \;\; \forall x = t-m,\dots,t-1, \;\; \mathbf{SST}(x) = 0 \\ 2 \;\; if \;\; \forall x = t-m,\dots,t-1, \;\; \mathbf{SST}(x) \geq 0 \\ -1 \;\; if \;\; \forall x = t-m,\dots,t-1, \;\; \mathbf{SST}(x) \leq 0 \\ 0 \;\; otherwise \end{cases} \tag{2}$$

In cases when $p_t \neq 0$, we say that it is possible to make a prediction, as there exists a trend, otherwise it is not, as the trend is ambiguous.

Simulation test of the proposed approach

In order to justify usage of the previously described signal strength prediction model, a number of simulation tests were conducted.

Simulation setup: Each simulation run consists of 20 mobile nodes moving following a modification of the Random Waypoint Model (RWP) [7] for 1000 steps. Each node is measuring signal strength from every other node, and predicts signal strength using (2) for each discrete time point. These predictions are compared with the actual measured values for each point, giving an evaluation of the prediction algorithm. Both the *prediction possibility* and the *prediction accuracy* are evaluated. *Prediction possibility* represents the average portion of simulation steps in which the value $p(t)$ from (2) is nonzero, i.e. a prediction is possible. *Prediction accuracy* represents the average number of correct predictions a node made throughout the simulation.

Random waypoint model: Nodes are assigned random starting location within a 100×100 square. Each node's movement pattern consists of straight paths of different step lengths, and each path has it's own speed and angle. Speed for each path is uniformly distributed in $[1, 3]$ m/s range, path lengths are uniformly distributed in the following ranges of steps: $[1, 5]$, $[3, 10]$, $[5, 15]$.

[1]This is a consequence of the underlying idea of alternative topology creation algorithm. In this algorithm, a stable signal strength is regarded as a good thing.

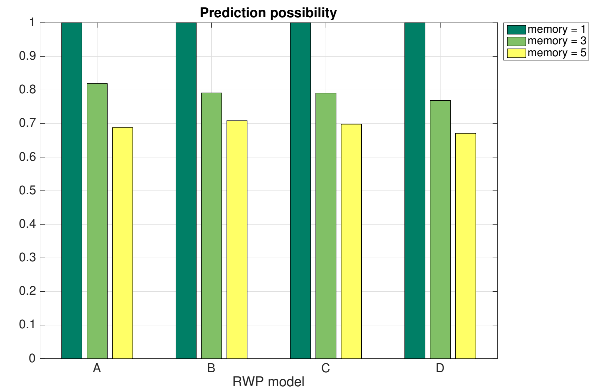

Fig. 2. Possibility to predict for different mobility models

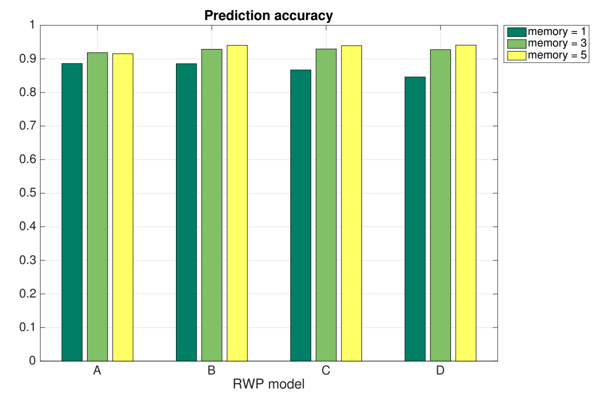

Fig. 3. Prediction accuracy for different mobility models

We approach the angle parameter in two different ways. The first one assumes an independent angle for each path, with uniform distribution, and the second one takes the *change* of angle into consideration. This way, we are trying to make the mobility model more "human", limiting the angle difference from one path to the next to better resemble actual human movement.

Angle parameter for different mobility models used in different simulation runs are listed below, and figure 1 shows node paths in each of the different scenarios:

- **A** – Angle distribution $Unif \sim [-180, 180)$
- **B** – Angle change distribution $Unif \sim [-90, 90]$
- **C** – Angle change magnitude distribution $Exp \sim [0, 180]$, sign $Unif \sim \{-1, 1\}$
- **D** – Angle change magnitude distribution $Exp \sim [0, 90]$, sign $Unif \sim \{-1, 1\}$

Simulation results: Different combinations of simulation parameters were used to test the performance of the proposed prediction technique. The general conclusion for all the setups is that there exists a trade-off between the *prediction possibility* and the *prediction accuracy*, and is influenced by the memory size. The shorter the memory, the greater is the probability of a clear trend appearing, while the longer memory ensures more precise recognition of patterns.

There is also a link between the memory size and path length ranges, in a sense that if memory length falls into the path length range, it expectedly degrades the performance. For path lengths longer that the length of memory, which is a realistic scenario, we get more stable results. In general, shorter path lengths are less predictable, but are also a less realistic representation of human movement, so we opted for the longest range for our simulations.

Simulation results for different mobility models, and different memory sizes are shown on figure 2 for *prediction possibility*, and figure 3 for *prediction accuracy*. The step range for mobility paths used is [5,15]. Moving average is applied in order to reduce fading impact.

As stated above, the trade-off between the *prediction possibility* and the *prediction accuracy* is obvious. This is because a longer memory means that a trend needs to exist for a longer period of time, which allows for a more stable detection of trends, and a shorter memory means it is easier for trends to appear, but they are less stable. This allows for setting the memory size parameter based on the actual needs of the system using this technique.

This simulation test shows the justification of our approach to signal strength prediction. This approach, while admittedly simple and purely intuitive and empirical, shows potential for even better performance with some additional improvements dependant on the system it is going to be used in.

IV. Cooperative Data Distribution Application

As described in Section II, the existing cooperative data distribution solutions all suffer from the inability to detect and anticipate node mobility and its effects. Being able to predict, with a certain probability, the changes happening in the network in the near future can improve the network performance. Following this idea, we can design an algorithm to create an alternative network topology. This is a database formed by each node, which does not reflect the logical or physical topology, but rather the suitability of each neighbouring node for cooperation in data distribution. The first step we are implementing towards creating a full alternative topology, described in Section I, is signal strength prediction.

Network model

The idea of cooperative data distribution mechanism using caching assumes that nodes cache some data in order to serve as a source for another node that might request the same data later on. Figure 4 shows a common scenario where, due to mobility of the nodes, a connection is lost and another one needs to be established in order for the file transfer to be continued. Yellow node is the requesting one, blue nodes do not have the data cached, and red nodes do. First, a request is made and a source is chosen. After nodes start moving away, the connection is broken, and a new source is needed.

In order to test the application of the described signal strength prediction algorithm, we decided to apply it to a generalised mobile ad hoc data distribution network testbed. Mobile nodes are connected in an ad hoc network. One node is requesting data – *client*, and the others act as data *sources*.

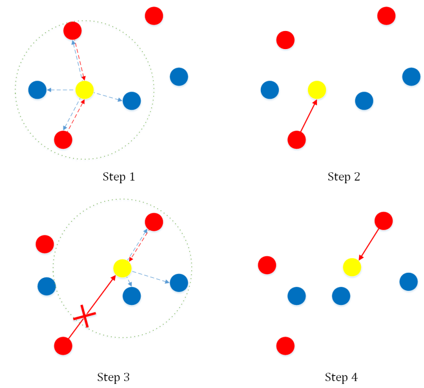

Fig. 4. Source change due to node mobility

Data is represented by a file split to equal sized numbered pieces (100000 of them, each 1024 bytes large), which are individually transmitted. Each source is considered to have all the pieces.

Client initiates the data exchange by sending a unicast request for the first piece, to the selected source node. In case of an interrupted transmission (for whatever reason – different scenarios described later), the following data request from client includes the number of the first data piece not yet received by the requesting node. Sources send data to the requesting node until the connection is lost or stopped by the client. Every data transmission is timed by the client, from the first request, until the last piece is received.

Alternative topology creation

The client node is maintaining its alternative topology containing data for all the other nodes in the network. Based on (2), each of the other nodes is assigned a prediction value. This value represents the likelihood of the node's signal strength to increase or remain the same. The predicted values are accumulated over five steps and represent an *evaluation* of the signal strength trend. This trend, combined with the current signal strength, makes up the final score for each of the nodes' *suitability* as a data source from client's perspective.

Different combinations of the two parameters, *signal strength*, and *trend evaluation*, have been tested, and the current experimental testbed setup showed that using a simple sum to give the best results in the sense of stability of source choice and the ability to correctly detect network changes.

Signal strength prediction mechanism parameters were also varied, and we have empirically reached a conclusion about the values that best fit the described testbed network.

Following equations show the workings of the alternative topology creation algorithm, where s is *suitability* – the measure of predicted quality of a node as a data source, e is *evaluation* – the accumulated prediction values over time, m_{eval} is the number of previous signal strength predictions taken into account, ss is signal strength, p is predicted value and is calculated as in (2), with the measuring and prediction interval being (Δt). All the values are for a single neighbouring node.

$$e(t) = \sum_{i=1}^{m_{eval}} p(t - i * \Delta t) \tag{3}$$

$$s(t) = ss(t) + e(t) \tag{4}$$

The node with the highest *suitability* value is chosen as the preferred source. In case this is not the currently active source, a source change is invoked, and the data transfer continues from the new source. The transfer is continued from the same point where it stopped from the old source. A back-off interval is applied in order to prevent too frequent source changes, known as the *ping-pong effect* [8]. Another restriction put upon a source change is that it can only be performed when the active transfer rate falls below a set threshold value.

Test scenarios

Three different scenarios are tested. The *default/passive*, *reactive* and *proactive* one.

1) Default/Passive: In this scenario, client scans signal strengths from neighbouring nodes, and chooses the one with the strongest signal as the data source. It sends the chosen source initial data request, and starts receiving data pieces. This transmission is active until the link breaks down, after which the client repeats the same process – scans for the strongest signal, chooses the new data source, and sends a request to it, from the next data piece in line to be received. This is a *passive* approach, where absolutely no awareness of the network surroundings is applied, nodes only detect an interrupted transmission, and try again.

2) Reactive: In this scenario, client maintains the alternative topology, in its current form (only signal strength prediction). Signal strength and its trend are followed and predicted as described in Section II. Based on this evaluation client chooses the initial data source, and sends the data request. Monitoring signal strength is continued throughout the whole time of data transmission, but this knowledge is not used until the active link breaks down. As soon as that happens, the new source is selected by its suitability parameter, reflecting the signal strength trend, and the request for the next piece is sent. This approach utilises the benefits of network awareness, but only after there is an interrupted transmission.

3) Proactive: This scenario is similar to the previous one, with one important difference. Alternative topology measures are not only maintained throughout the whole duration of data transfer, but are constantly monitored in order to choose a better data source. This means that if the active transmission link starts deteriorating, a more suitable source, if there is one, will immediately be chosen as the new data source, the active transmission will be stopped, and the new one initiated. This way, a higher data transmission rate is preserved, allowing for a faster data transfer time.

Experimental testbed setup

Our testbed consists of six Raspberry Pi 2 model B [9] computers (nodes), running a customised version of Linux operating system – Raspbian [10]. Each node is equipped with

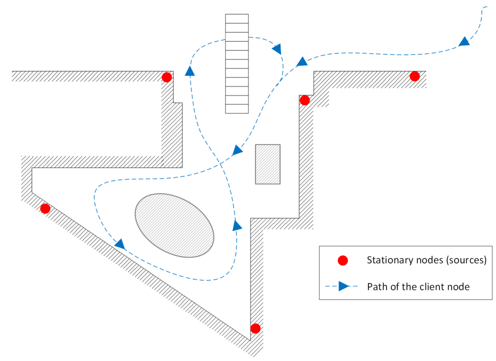

Fig. 5. The testbed plan

a USB WiFi adapter, supporting 802.11bgn standards [11], and is powered by an external battery, allowing full mobility and autonomy.

The data distribution algorithms for the previously described scenarios are implemented using Python [12] scripts running on each device. Each scenario was tested on ten runs, with the client node following the same path every time.

Figure 5 shows the plan of our experimental testbed, as well as the locations of the static nodes, and the path of the mobile node (client). The testbed is located in a garden with approximate dimensions of around 30×35 m, which allows for nodes to be positioned in a realistic way and for their ranges to overlap significantly, making it a good scenario to test the proposed approach.

Results

In this section we will present and analyse the results of the previously described testbed experiments.

The results show a significant difference in file transfer time between the proactive approach, and the other two. Figure 6 shows the average data transfer times for each of the three scenarios. The proactive approach has the best file transfer speed, resulting in an average transfer time of around 58 seconds, compared to the ones in reactive, 100 seconds, and passive, 141 seconds.

As all of the experiments have been conducted following the same path and with the same locations of the nodes, there is an obvious overlap of the three scenario's results in the first 5 seconds of the experiment. This is due to the client node always detecting the same node as the source at the start.

What is also noticeable, is that while the file transfer speed starts decreasing (the slope of the graph gets flatter) for the other two scenarios, the proactive one maintains a relatively constant rate. This happens every time the client node moves away from the active source, and the implemented alternative topology algorithm allows it to choose a better one.

In both reactive and passive scenarios the client node waits for the connection to the active source to break before requesting the remaining data from a new source. The way they choose this new source differs, and is responsible for the

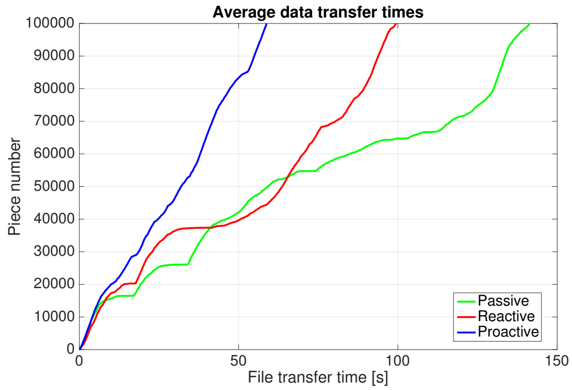

Fig. 6. Average data transfer times for the three scenarios

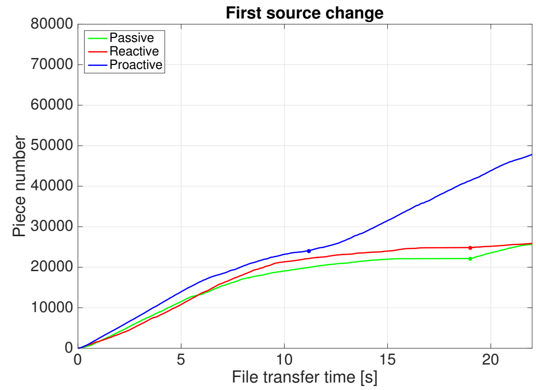

Fig. 7. First source change for one run of each of the scenarios

better results achieved in the reactive scenario. As described before, the client in passive scenario chooses the new source based only on the signal strength measure, while in the reactive one, it is using our alternative topology algorithm, making a decision based on node mobility perception through trends of signal strength measures. As a consequence, the client node made fewer source changes, but better ones, as the overall file transfer time was shorter then in the passive scenario, resulting in 3.6 source changes on average in the passive scenario, and 3.1 in the reactive one. In contrast, the client in the proactive scenario made 5.6 source changes on average, but those were always induced by a better source being available, which explains the almost constant data transfer speed.

Figure 7 is an example of the described differences in all three scenarios' behaviour. It shows how, in the same conditions, the client changes source in the proactive scenario much sooner compared to the other two scenarios (dots on each line show the moment of source change). It does so as soon as the data transfer speed starts decreasing, and a new, better source is available. The example in the figure is from the beginning of the file transfer, when in all of the three scenarios the same initial source is chosen, and after starting to move away from it, the client changes the active source at different times.

These results show that our approach allows nodes to predict the deteriorating and connections increasing in quality, and using this knowledge switch between the sources, maintaining a high data rate throughout the entire file transfer.

V. Conclusion

In this paper we presented our idea of an *alternative topology*, a probabilistic measure of neighbouring node suitability for cooperative data distribution in MANETs. The algorithm for the formation of this measure is currently based on predicting future signal strength based on its trends. We verified our approach using a simulation, in order to apply in on an experimental testbed for MANET cooperative data distribution.

Results of our file transfer experiment show that using our algorithm, a significant reduction in file transfer time can be achieved compared to a passive approach, as a relatively constant file transfer speed is maintained, thanks to predicting the relation of nodes and always choosing the best source available.

We are currently building on the work presented in this paper, as our *alternative topology* should also consider other parameters, as speed and direction of movement of the nodes, node's neighbourhood stability and cache content. This should increase the level of awareness nodes have about the network dynamics and further improve data distribution performance.

References

[1] B. Liu, V. Firoiu, J. Kurose, M. Leung, and S. Nanda, "Capacity of cache enabled content distribution wireless ad hoc networks," *Proceedings - 11th IEEE International Conference on Mobile Ad Hoc and Sensor Systems, MASS 2014*, pp. 309–317, 2015.

[2] Y. Du, S. K. S. Gupta, and G. Varsamopoulos, "Improving on-demand data access efficiency in MANETs with cooperative caching," *Ad Hoc Networks*, vol. 7, no. 3, pp. 579–598, 2009. [Online]. Available: http://dx.doi.org/10.1016/j.adhoc.2008.07.007

[3] H. Artail, H. Safa, K. Mershad, Z. Abou-Atme, and N. Sulieman, "COACS: A cooperative and adaptive caching system for MANETs," *IEEE Transactions on Mobile Computing*, vol. 7, no. 8, pp. 961–977, 2008.

[4] S. Lim, W. C. Lee, G. Cao, and C. R. Das, "A novel caching scheme for improving Internet-based mobile ad hoc networks performance," *Ad Hoc Networks*, vol. 4, no. 2, pp. 225–239, 2006.

[5] L. Yin and G. Cao, "Supporting Cooperative Caching in Ad Hoc Networks," *IEEE Trans. on Mob. Comp.*, vol. 5, no. 1, pp. 77–89, 2006.

[6] K. Farkas, T. Hossmann, F. Legendre, B. Plattner, and S. K. Das, "Link quality prediction in mesh networks," *Computer Communications*, vol. 31, no. 8, pp. 1497–1512, 2008.

[7] D. B. Johnson and D. A. Maltz, "Dynamic Source Routing in Ad Hoc Wireless Networks," *Mobile Computing*, vol. 353, pp. 153–181, 1996. [Online]. Available: http://www.springerlink.com/index/10.1007/b102605

[8] T. Inzerilli, A. M. Vegni, A. Neri, and R. Cusani, "A location-based vertical handover algorithm for limitation of the ping-pong effect," *Proceedings - 4th IEEE International Conference on Wireless and Mobile Computing, Networking and Communication, WiMob 2008*, pp. 385–389, 2008.

[9] "Raspberry Pi." [Online]. Available: https://www.raspberrypi.org

[10] "Raspbian." [Online]. Available: https://www.raspbian.org

[11] L. A. N. Man, S. Committee, and I. Computer, *Part 11 : Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) S pecifications IEEE Computer Society*, 2012, vol. 2012, no. March.

[12] "Python." [Online]. Available: https://www.python.org