

INTERNATIONAL JOURNAL OF COMPUTERS COMMUNICATIONS & CONTROL  
ISSN 1841-9836, 13(4), 550-565, August 2018.

# Improving DTNs Performance by Reduction of Bundles Redundancy using Clustering Algorithm

R.O. Schoeneich, P. Prus

Radosław Olgierd Schoeneich\*, Piotr Prus

Institute of Telecommunications

Warsaw University of Technology, Poland

\*Corresponding author: [rschoeneich@tele.pw.edu.pl](mailto:rschoeneich@tele.pw.edu.pl)

[p.prus@stud.elka.pw.edu.pl](mailto:p.prus@stud.elka.pw.edu.pl)

**Abstract:** This article presents complex clustering algorithm for Delay Tolerant Networks (DTNs) including neighborhood discovery, cluster creation, and data distribution within cluster. The general idea is to reduce the amount of messages being sent and of buffer utilization, by taking advantage of the nodes tendency to create groups and to share similar mobility patterns among each other. The main purpose of the algorithm is to improve the network performance without major changes in communication schemes between nodes. Almost no extra message type is added. Most extra features are available thanks to adding small extra fields into transmitted packages. Neighborhood discovery is being realized passively by listening to other nodes messages. The proposed algorithms allow to reduce both the bandwidth occupation, as well as the problems related to the media access. Furthermore, it can increase message delivery probability thanks to intelligent package distribution inside created cluster. Simulations were carried out to evaluate the effectiveness of the proposed solution in terms of package delivery probability, mean buffer occupancy and mean hop number to delivery the message. Results of simulation show that this solution is not necessary or recommended for small-scale networks with few nodes using clustering algorithms. However, with increasing number of nodes and messages, the performance of non-clustered DTNs drops significantly while clustered network works efficiently.

**Keywords:** Delay Tolerant Networks (DTNs), mobile ad-hoc networks, clustering algorithms, cluster based routing, neighborhood discovery in ad-hoc networks.

## 1 Introduction

Delay Tolerant Networks (DTN) [2] [9] are a kind of networks that can be used in extreme terrestrial environments with no telecommunication infrastructure provided, e.g. deserts, or rain forests. This kind of network can also be used in emergency situations, as an alternative way of communication, when conventional networks break down, for example as additional communication channel for emergency services in the case of terrorist attack. It can be also used for military purposes, as communication channel between groups of soldiers. The main advantage of this kind of network is the fact that no infrastructure is needed for its operation. Therefore, it can be easily created in any situation. Family of DTN networks that is taken into account is mobile ad-hoc networks that may lack continuous connectivity. It relates to situations when, in a particular moment, direct path between two nodes may be not available. However, communication is done through message ferring by mobile nodes. DTN are rather low transfer rate networks. The main assumption of DTN is to deliver the particular message as fast as possible, with maximum probability of delivery. However, in DTN networks it is possible that the path between two nodes will never appear, so the message might not be delivered.

The main mechanism of message forwarding is presented in Figure 1. A node that has a message to send broadcasts INV - the Invite message. This message contains a list of content identifiers, which node stores in its buffer. When a node that has received the INV message

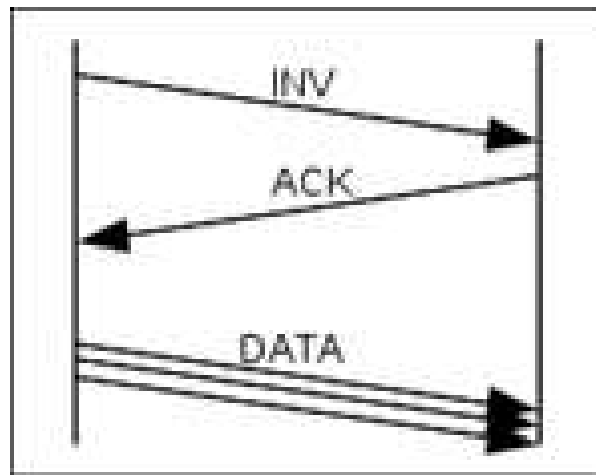


Figure 1: Basic message exchange scheme for DTN networks

wants to get particular packages, it sends ACK - the acknowledge message. The ACK message contains the list of desired content ids. After receiving ACK message, the first node forwards desired packages via DATA packages. This mechanism is being repeated until message reaches its final destination. INV messages are being sent periodically. Routing protocol that uses this mechanism is called Epidemic Routing [32]. It is a flooding protocol. It means that the message is being sent to each node met. Epidemic routing allows to reach highest possible delivery probability, along with the shortest delivery time in particular mobility pattern.

The main problem of this protocol is high redundancy. With increasing number of packages, the probability of message delivery drops significantly, because of limited buffers capacity. The second important problem is connected to the media access control. When nodes travel in a group, each node buffer will contain exactly the same content after some time. This means that each node will send identical INV messages. With increasing number of nodes, the access to the radio channel becomes more difficult. There is no need for each node of the group to send the new additional INV message. More advanced DTN routing protocols, such as Spray and Wait [29] or PRoPHET [20], limit the redundancy by reducing the number of message copies forwarded. On the one hand, this solution reduces the utilization of buffers. However, on the other hand, it removes some possible paths, which may lead the message to the destination. When mobility pattern is not completely random, two nodes will meet each other sooner or later, and the delivery probability drops. In some mobility patterns, in which nodes are constantly mobile, there are many possible paths from the source to the destination. In such patterns, the number of copies forwarded can be reduced without a significant performance loss. On the other hand, for the mobility patterns, in which only some nodes are mobile, or nodes are moving in specified directions, the number of possible paths is much lower. Reducing number of forwarded copies in such networks is highly ineffective. Moreover, advanced routing protocols do not solve the problem with radio channel access.

This considerations show two serious problems considering scalability of DTN networks: buffer and radio channel capacity. Wise use of clustering algorithms can solve both problems. Data distribution within improves significantly the message capacity and reduces the number of transmissions of the INV message, which increases the medium usage efficiency.

The idea of clustering has been extensively investigated in ad-hoc networks. First solutions like Clusterhead Gateway Switch Routing [5] and Cluster Based Routing Protocol with extensions [28] [1] [31] was reference for next development. All of these solutions introduce the idea of two

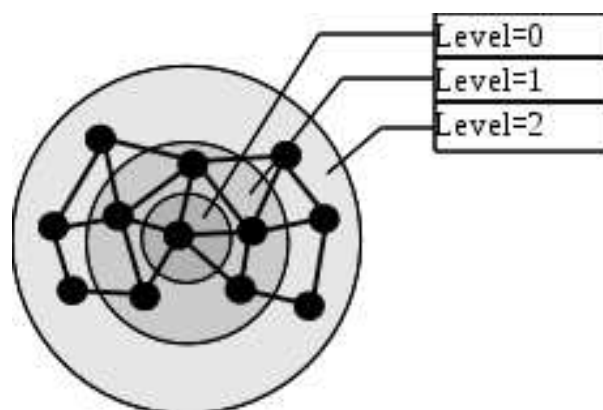


Figure 2: Two level cluster topology

level hierarchy composed of higher level managing nodes - Cluster Heads and lower-level nodes - Cluster Members with one-hop single cluster dimension. The multi-level hierarchy was proposed in [14]. Multi hop cluster range was proposed in [34].

For WSN networks there were proposed similar solutions: based on two-level single hop clusters like Low-energy Adaptive Clustering Hierarchy [13], multi hop clusters e.g. Hybrid Energy-Efficient Distributed clustering [36], or multi-level clusters like in Distributed Weight-based Energy-efficient Hierarchical Clustering protocol [8] and Two-Level Hierarchy LEACH [24]. The other very popular routing protocols which are based on clusters in sensor networks are: PANEL [3], UCS [30], EECS [35], EEUC [19], ACE [4], BCDCP [27], PEGASIS [21], TEEN [25], APTEEN [26], TTDD [23], CCS [16], HGMR [18], solutions like [10] [11] [12]. The detailed survey of clustering algorithms can be found in [33] and [22].

The clustering in DTN routing application was discussed in DTMN protocol [7] where authors proposed the solution for grouping mobile nodes with similar mobility pattern into a clusters. Similar solution for nodes clustering in vehicle-based networks was proposed in [6]. Other solution called hierarchical clustering-based forwarding scheme (HCS) was proposed in work [17]. The solution is based on implementing hierarchical clustering on social information.

All these works are focused on routing schemes in wireless networks, the message storage efficiency is less important. Our work focuses on achieving a high performance of the message storage and on the elimination of redundancy in a complex DTN network environment.

## 2 Proposed clustering algorithm

### 2.1 Cluster topology

Proposed algorithm involves grouping nodes with similar mobility patterns into two-level clusters. The general scheme of such cluster is presented in Figure 2.

A node in the area called Level=0 is called Cluster Head (CH). This node is responsible for cluster management. Address of this node is also considered as cluster id. Level=1 nodes are directly connected to CH. These nodes are responsible for data storage. Level 2 nodes are regular cluster members. The decision to organize the nodes into two-level clusters is due to the fact that such structures are still quite easy to manage and maintain in terms of their consistency. Taking into account the range of radio communication, such approach is enough to cover area of node group with a similar mobility pattern. The structure of INV message sent by DTN network with cluster algorithms is presented in Figure 3.



Figure 3: INV message structure

The package contains information about address of message sender, cluster id, and its position in cluster. DATA part of message stores a list of packages available to be sent. When a node does not belong to any cluster, the value 0 as CH address is set.

## 2.2 Related nodes metric

The main problem in DTN Networks is the fact that nodes are constantly mobile, which makes the cluster creation and maintenance tasks difficult. Network topology changes constantly and, furthermore, the connection between two nodes can disappear permanently. We assume that nodes do not know their position via GPS etc. The only information about the neighborhood we acquire is by listening to other nodes messages. To determine whether two nodes belong to the same cluster, some additional mechanisms have to be implemented. The solution to this problem has been presented on the basis of the Clustering and Cluster-Based Routing Protocol for Delay-Tolerant Mobile Networks [7]. Because of the fact that each node sends INV messages, there is no need to send extra HELLO messages, which are common for typical cluster-based ad-hoc networks. All the necessary/required information can be acquired from periodically sent INV messages. The period between INV messages is known. For each time slot, the related nodes metric is being evaluated. In general form, the metric formula is as follows:

$$RN_y[n] = x_y[n] + \alpha^1 x_y[n - 1] + \dots + \alpha^N x_y[n - N]$$

Where:

$\alpha$  - the extinction coefficient,  $y$  - the identifier of neighbouring node,  $x_y[n]$  - the binary factor which indicates if node received message from node  $y$  in time slot number  $n$ ,  $N$  - number of back slots,  $RN_y[n]$  - value of the metric.

By changing the metric parameters, the different behavior of nodes is being set. Parameter  $\alpha$  determines how long the information about relation is important. The higher the value of  $\alpha$  the smaller impact on the metric value the previous meeting have. Parameter  $N$  determines the number of time slots for which the metric is being evaluated. On the one hand the higher  $N$ , the more reliable the metric value is. On the other hand, it increases the memory requirements. Additionally, the information about meetings from distant past is useless, because of a dynamic network topology change. Value of parameters have to be set individually for each situation. Depending on evaluated metric, two thresholds are defined, i.e.  $Y_{tresh}$  and  $Y_{min}$ . These thresholds will be used to make decision. If the condition:

$$RN_y[n] > Y_{tresh}$$

is fulfilled, then we assume the node is strongly related with node  $y$ . On the other hand, if the condition:

$$RN_y[n] < Y_{min}$$

is fulfilled, the relation between two nodes broke, and they no longer can be considered as strongly related nodes. These simple calculations allow to distinguish nodes with similar mobility pattern, which is very important in cluster formation and management. Metric also stores information about cluster assignment of observed node.

## 2.3 New cluster formation

If the number of the strongly related neighbors, which also do not belong to any cluster, exceeds threshold, the procedure of cluster initialization is carried out. Node broadcasts OR-

GANIZE message, which contains address of a sender and number of strongly related neighbors. Each node, which receives such message has to answer and sends its own ORGANIZE message. Node with the highest number of neighbors becomes a new Cluster Head, and its neighbors become Level 1 Cluster members. Such approach can lead to creation of two clusters, if two nodes are not in radio range, however, other mechanisms, which will be described later, will adjust this situation.

## 2.4 Joining new node to cluster

In order to join cluster, several conditions have to be fulfilled. Each node listens to broadcasted messages and evaluates related nodes metric. Nodes have to be strongly related to a sufficient number of cluster members to be concerned as strongly related to cluster. Cluster topology provides two levels, so metric is only valid for Level 0 and Level 1 nodes. If node determines that it is strongly related to the cluster, it can send a message requesting to join the cluster. Node becomes a part of the cluster of an adequate level. All packages from its buffer are being sent to the cluster and then distributed among Level 1 cluster members.

## 2.5 Leaving the cluster

Because of node mobility, cluster consistency has to be managed constantly. Node located at Level 1 has to be strongly related to Cluster Head, and a certain number of Level 1 nodes. If the first condition is not fulfilled, node changes its Level to 2. If second condition is not satisfied, node must leave the cluster. For Level 2 nodes only the second condition has to be fulfilled. Leaving node sends appropriate message to the Cluster Head and detaches from the cluster. It takes copies of as many packages from the cluster as its buffer can handle.

## 2.6 Two clusters meeting

Situation when two clusters are within radio communication range is hard to manage. CH reelection and cluster reorganization common for cluster-based ad-hoc networks will be highly infective due to packages stored inside cluster. This could lead to the package loss caused by limited buffers capacity, and high network load associated with the package transmission to new Level 1 nodes. Proposed architecture prefers absorption of smaller clusters by the bigger ones. If node determines that it is better related to an adjacent cluster, it will become a member of a new cluster. However, it is required to counteract the situation, when node oscillates between two clusters. This can be achieved by defining another threshold. Node can change cluster only if relation with new cluster is greater than current increased by threshold. This approach leads to the absorption of less numerous clusters by bigger ones.

## 2.7 Cluster destruction or cluster head reelection

Last node leaving the cluster is the Cluster Head node. When nodes are leaving cluster, number of strongly related nodes seen by the Cluster Head is decreasing. When the number drops below a particular value the Cluster Head makes decision to destroy the cluster and broadcasts DESTROY message. Reaction to such event can be different depending on the situation. Decreasing number of strongly related nodes seen by CH may be result of Cluster Head mobility pattern change leads to leaving the group. Node which receives DESTROY message checks a number of its strongly related neighbors. If it exceeds threshold of a new cluster creation, new Cluster Head is being reelected as in new cluster formation algorithm. If

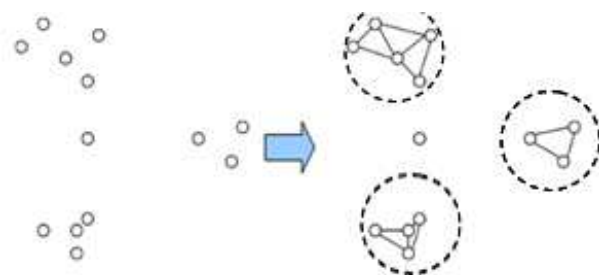


Figure 4: The network topology after cluster formation

the number of strongly related nodes is too small, and no ORGANIZE message was heard by different node, the cluster is destroyed.

## 2.8 Cluster head loss

Architecture must be resistant to the Cluster Head node failure. Without additional mechanisms, such situation would lead to continuous migrating nodes to Level 2 area and eventually to destruction of the cluster. To prevent the situation when Level 1 node loses a strong relation to the Cluster Head, the first node ensures that the neighboring nodes see the Cluster Head node. A broadcast message is sent and response is expected. If message is not received, the Cluster Head re-election occurs. New Cluster Head is reelected among Level 1 nodes. This makes the reorganization of cluster much easier because Level 1 nodes are better related and closer to the disappeared Cluster Head.

## 3 Cluster based routing

Organization of clusters changes topology of DTN networks. Nodes with similar mobility pattern are grouped into clusters creating bigger virtual nodes of network. General idea is to treat clusters like regular DTN nodes. These nodes share functionality of storing and forwarding messages, however, thanks to organization, their buffer capacity is much bigger. In clustered DTN, the number of INV messages sent by the cluster is smaller than in flat networks. This makes clustered networks more scalable than regular DTN networks. Network topology after cluster formation is presented in Figure 4.

### 3.1 Inter-cluster routing

Creation of virtual nodes allows coexistence of clusters and regular DTN nodes. No additional mechanisms have to be implemented. Message exchange between two nodes is based on Epidemic Routing protocol described in introduction. The choice of this protocol is associated with its high performance under condition of package delivery probability and delivery time. Furthermore, this protocol is easy to implement and do not need any additional information storage or evaluation. Problem with radio channel access control is solved by preventing some nodes to send INV message. The solution is to introduce the probability of INV message sent. Nodes with more neighbors send INV messages less often than these with lower number of neighbors. When randomly generated number is higher than threshold, the INV message is sent. Value of threshold is evaluated based on neighbors' quantity. After leaving too many time slots, the empty INV message is sent to indicate that node is still inside cluster. This approach limits the number of INV message sent without significant performance loss of message delivery. It is important to

mention that the Cluster Head has to broadcast INV messages in each time slot due to specific function of this node in cluster management. Lack of connection with Cluster Head has to be counteracted as fast as possible.

### 3.2 Intra-cluster routing

Nodes grouped into cluster creates consistent ad-hoc network. It is organized into two level hierarchy with one central node. Moreover nodes have information about their neighbors based on related nodes metric evaluation. This makes creation of routing tables inside cluster easy. Most appropriate routing protocol in such structure is Optimized Link State Routing Protocol (OLSR) [15] typical for ad-hoc consistent mobile networks. Path between two nodes within a cluster is almost always known. Each node know which neighbor is connected directly to the Cluster Head. Transmissions inside clusters are made on demand. Data transfer inside cluster is much faster than between two DTN nodes. There is no offset made by time slots typical for INV messages.

## 4 Data handling

An objective of the proposed method was to reduce the redundancy of packages in DTN networks. It is also required that network load reduction does not cause performance degradation. Single node failure, or unexpected cluster leaving, cannot lead to data loss. The purpose of this section is to describe the specifics of data handling for DTN cluster-based network. It presents the general assumptions and methods of data distribution within a cluster.

Due to their specific nature, no message delivery confirmation methods are implemented in DTN networks. Using this methods would result in message flooding in the opposite direction. That means an additional load on the network without noticeable benefit. Common concept is using Time To Live (TTL) parameter. After the time expiration, message is removed from the buffer. This solution has its explanation. If any node demands message for a long time, it may be assumed that it has already been duplicated and forwarded. Optimal value of TTL is specific for each network, and should be set individually.

### 4.1 New package distribution in cluster

New message in the cluster can only occur as a result of receiving the INV message from encountered node. The process of adding new nodes to a cluster takes place over several INV messages time slots and there is no need to consider the situation in which a new attached node carries packages that are not already in the cluster. Each new package received by the cluster is sent to the Cluster Head node. Then the message is forwarded to Level 1 nodes. To avoid unexpected losses associated with unexpected node failure, the message is replicated in a number of nodes simultaneously. The Cluster Head remembers package sequence number and information about nodes that possess it. The Cluster Head manages the packages and in case of node failure sends the requests to replicate the package in another Level 1 node. The CH manages the packages to be evenly distributed in nodes. The nodes through which a package has passed stores it through the entire duration of the TTL. This limits the number of transmissions and creates cache like buffers. After placing message in buffers the Cluster Head broadcasts the message about new available content.

## 4.2 Data search

INV messages sent by cluster members contain a list of all packages available. If ACK message is received the data search mechanism starts. First own buffer is being searched. If content is found, it is passed to requested node. If no data was found in buffer, request to the Cluster Head is sent. Level 2 nodes send request through Level 1 node. At first this node is also searching for its own buffer. If any of them finds desired content, The Cluster Head is asked. The CH responds with address of the node, which holds the content. Then package is downloaded directly from the node. Afterwards the package can be forwarded to desired location.

## 4.3 Buffer overflow

There may occur a situation, in which buffer of particular node overflows. Threshold of 0.95 buffer capacity is set. If the buffer occupation exceeds this number, some content has to be deleted. There are two types of package in buffers: created by the Cluster Head, and created when package was forwarded by the node. The first to delete are cached messages. The node removes as much packages as necessary to reduce buffer occupation below the threshold. Messages are deleted in order of time spent in buffer. If removal of cached packages is not enough to fulfill the buffer occupation condition, request to the Cluster Head is sent. At least two copies of one content have to be stored in the cluster to avoid data loss in case of node failure. If there are more copies, the Cluster Head permits to delete the message. Otherwise it checks if there is any other Level 1 node with free space in its buffer. If such node exists, the package is forwarded to this node. In other cases the oldest package is removed.

## 4.4 TTL expires

In contrast to the standard DTN network cluster member nodes do not remove the messages with expired TTL automatically. When the Cluster Head notices that package expired its TTL broadcasts the message to delete desired content. Each node, which received such message, deletes this package from its buffer. The list of available content sent in INV messages is updated.

# 5 Simulations

## 5.1 Mobility model

The research was carried out using specialized mobility model for emergency and rescue situations designed by authors. It is a mobility model proposed by us to describe the mobility pattern of nodes in an emergency situation. Scenario describes the situation when bomb explodes in the office. It takes into account different groups of people. The model describes mobility of injured civilians or people fleeing in panic. Behavior of different rescue groups was implemented: police, firefighters and emergency medical services. This model is a complex attempt to maintain the mapping of different groups behavior in life-threatening situations. We find reasonable application for clustering algorithms in such situations. Creating groups is common behavior for all living beings.

In general the scenario consists of 3 phases: stable phase, chaos and rescue operation. During the stable phase civilians are moving around the office. They do their job, go smoke a cigarette, or sit at their desk. This phase is not interesting from the research point of view, however, it is used to generate a random distribution of nodes during the explosion. Example of distribution of nodes in the stable phase is shown in Figure 5.



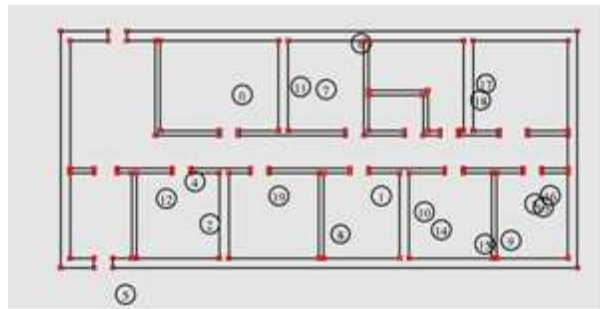


Figure 5: Node distribution during stable phase

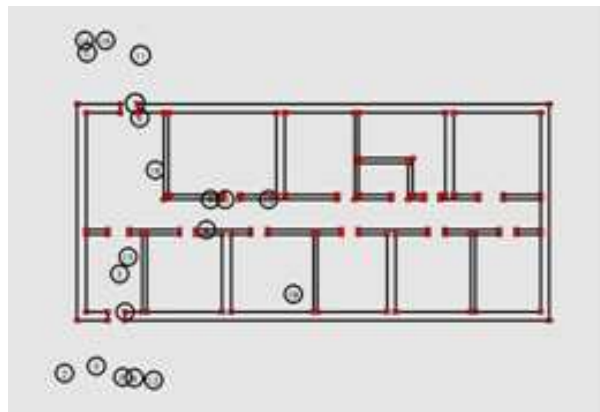


Figure 6: Node distribution during chaos phase

Stable phase is interrupted by explosion inside the office. Some civilians get injured, some are dead. Rest flees in panic. The phase of chaos starts. It lasts until arrival of rescue teams. Distribution of nodes in chaos phase is presented in Figure 6.

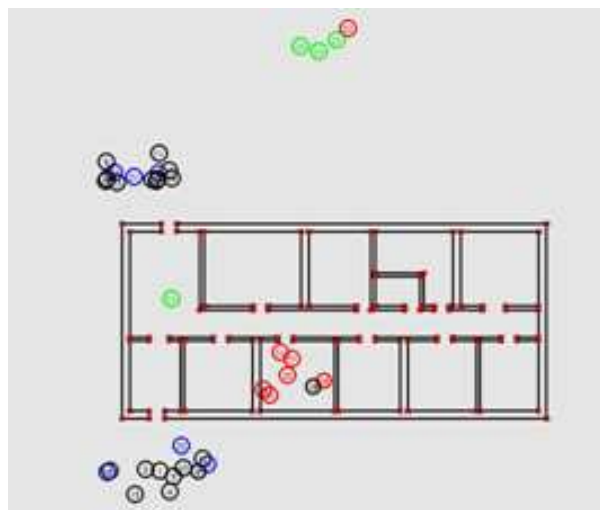


Figure 7: Node distribution during rescue phase

Evacuated people are grouping near office entrance. Here we can observe clusters. Those nodes are not highly mobile, but create a big buffer able to store message. Next, and last phase is the rescue operation. Firefighters enter the building, evacuate injured, police secure the area,

and separates onlookers. Medical services treat wounded people. Distribution of nodes in this phase is presented in Figure 7. Simulation ends when last wounded person is sent to the hospital.

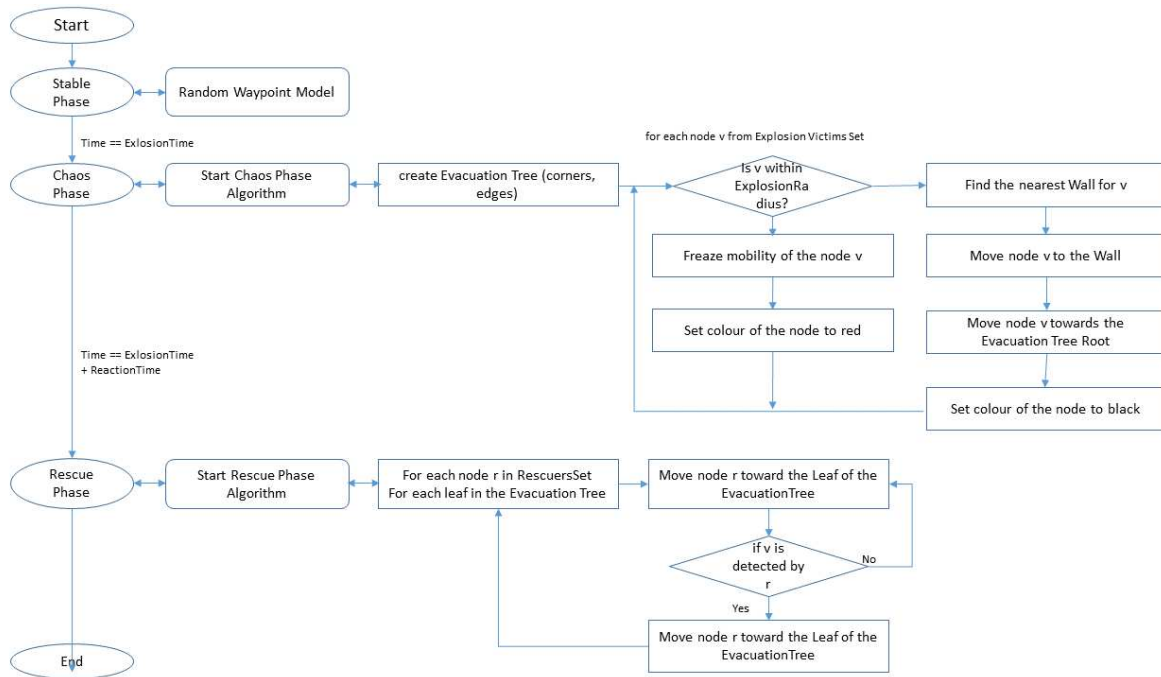


Figure 8: Mobility algorithm

The chaos phase has been implemented as follows: nodes inside a specific radius of explosion are seriously injured and are retained in place. These nodes are marked in red. Other nodes will escape using corridors out of the office. Due to smoke and reduced visibility nodes are moving along the walls. All mobile nodes in the phase are marked in black. To implement the mobility of nodes we proposed to use a set of edges of the door and edges of walls as points of the graph. All of this points are denoted as red dots. For each mobile node we executed algorithm that determines: 1) the closest distance between the black node and the nearest red point, then the algorithm 2) builds a graph between all red points and the building exit. 3) The algorithm performs the shortest path search in the graph based on the Dijkstra algorithm. After finding the shortest path, nodes are moved with speed of 2-4m/s to the exit. The rescue phase looks as follows: in simulations we propose that after a predefined period of time rescuers (green nodes) are moved. The movement is based on previous graph, but nodes are moved in opposite direction - to search all rooms in the office. Rescue nodes which hit a injured (red) nodes, take them to the exit. The mobility algorithm was presented in Figure 8.

## 5.2 Results

The research was performed in two stages. First for small scale network with small number of nodes and unlimited buffer capacity. This simulations show that clustered DTN works as efficient as non clustered network based on epidemic routing. The second stage is performed for a large scale network, where clustering algorithms show their advantage over regular DTN. During simulations following metrics were evaluated:

1. Package delivery probability: We counted the metric as number of messages received divided by number of messages send. The metric is important because in emergency and rescue scenarios many messages sent should reach destination nodes.

Table 1: Parameters of simulations

Parameter	Min Values	Max Values	Parameter	Value
$\alpha$	0	1	Radio Radius	5m
$N$	1	5	Bandwith	11 Mb/s
$Y_{tresh}$	1.45	1.45	Packet Size	1500 b
$Y_{min}$	0.95	0.95	Simulations Time	750 sec.
$Y_{cluster}$	3	30	MessagePeriod	180 sec.

2. Mean buffer occupancy: The metric describes the mean number of packets stored per number of nodes. It describes how efficient the solution is in terms of minimizing buffer occupancy and total redundancy.

3. Mean hop number: The Metric specifies how many nodes have to pass the message before it was delivered to the recipient. The purpose of this metric is to determine the efficiency of packets transmission and to estimate energy requirements. It also partly determines the efficiency of use of the capacity of the radio channel.

Proposed mobility model is random and after each run different result is generated. To obtain reliable result simulation have to be run several times. After 50 simulations the mean value, and confidence interval is evaluated. For small scale networks the conditions are as follows: 1) Unlimited buffer capacity, 2) Number of messages created: 1000, 3) Number of nodes: changes from 10 to 60

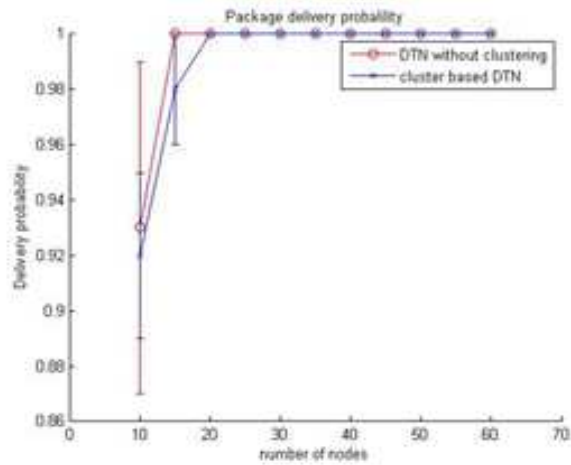
For all simulations we established nodes with different rules. Number of firefighters 6 and 1 officer, num of policemen - 6, num of rescuers - 4 persons. Total simulations time - 750sec., Messages sending period 180sec., The explosion moment was in 100 sec. after starting the simulation. Optimal values of parameters:  $Y_{tresh}$ ,  $Y_{min}$ ,  $Y_{cluster}$  and  $\alpha$  have been set based on preliminary simulations. The parameters of the algorithm are presented in the Table 1.

Results of simulations are presented in Figure 9 a, b, c.

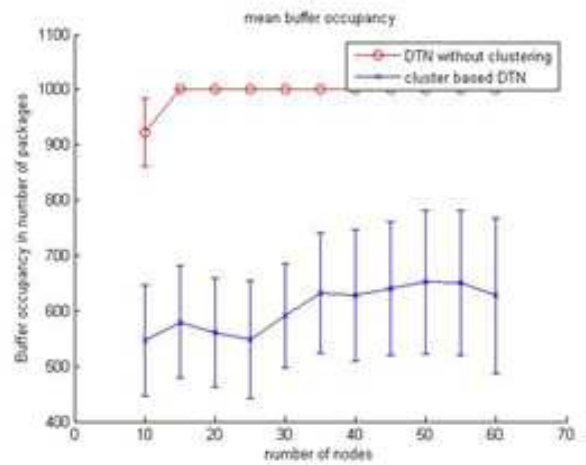
Results of simulation shows, that even in small scale networks cluster based DTN have advantage. Using clustering mechanisms do not reduce network performance within meaning of package delivery probability. Moreover it reduces the buffer occupancy and hop number. This leads to network load reduction. This advantages will be more noticeable in large scale network simulations. For those simulations conditions are: 1) Buffer capacity: 100 packages, 2) Number of messages created: 1000, 3) Number of nodes changes from 150 to 200.

Buffer capacity is smaller than number of messages. This should lead to overload of DTN network without cluster algorithms. Furthermore media access problems start being noticeable. Results of simulations are presented in Figures 10 a, b, c.

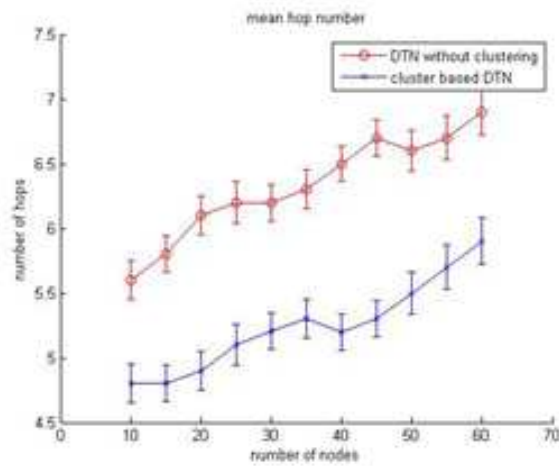
As expected regular DTN network performance drops significantly. When number of nodes exceeds 190 problems with media access are noticeable. Packages cannot be forwarded because network density is so high, that no INV message can be send without any collision. Message delivery probability is at unacceptable level. This is the result of buffers overflow. Hop numbers metric is evaluated only for successfully delivered messages, so this results do not carry any important information for regular DTN networks. More important information we get from results performed for clustered network. Hop number do not differ noticeably in comparison to small scale research. A slight increase in the number of hops is correlated with the increasing number of nodes in the network. This means that cluster based network forwards messages efficiently without additional, unnecessary transmissions.



(a) Package delivery probability

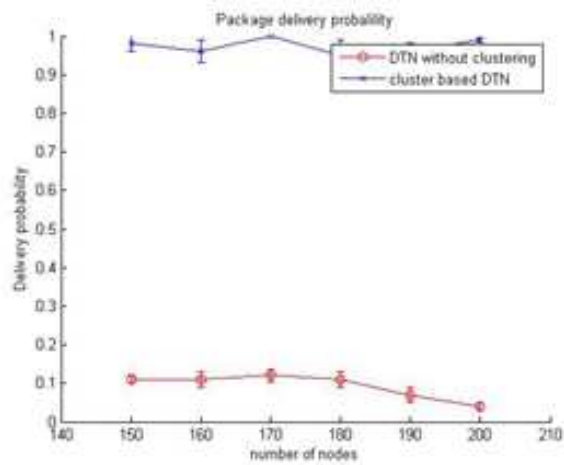


(b) Mean buffer occupancy

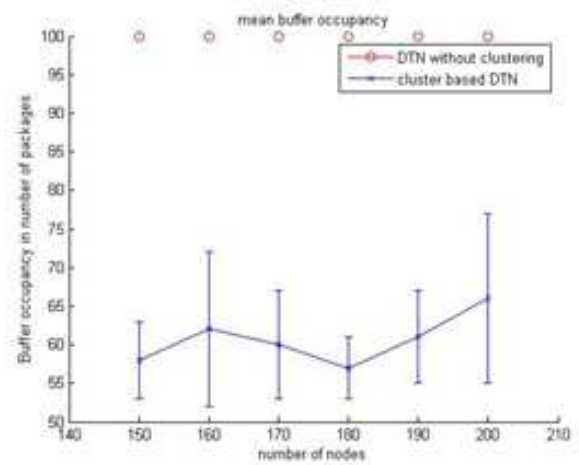


(c) Mean hop number

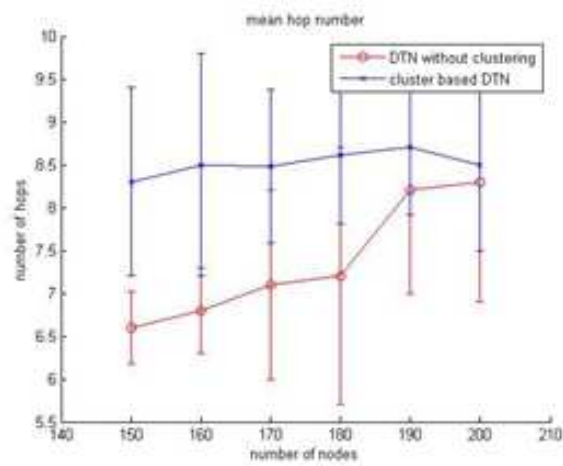
Figure 9: Results of simulations for small scale research



(a) Package delivery probability



(b) Mean buffer occupancy



(c) Mean hop number

Figure 10: Results of simulations for large scale research

## 6 Conclusions

We have proposed and implemented complex clustering algorithms for DTN networks. It significantly improves performance of such networks and makes them more scalable. Without many calculations, or sending extra control messages we achieved noticeable network load and buffer occupation reduction. This solution is designed to work in extreme terrestrial conditions where no telecommunication infrastructure is provided, and nodes moves in groups. This simple algorithm improves performance of regular DTN networks by reduction of unnecessary network load. Proposed solution can be easily implemented, and what is most important it can coexist with existing DTN networks

## Bibliography

- [1] Ahn, CW.; Ramakrishna, RS.; Kang CG. (2002); Efficient Clustering-based Routing Protocol in Mobile Ad-Hoc Networks, *in Proc. IEEE VTC 02*, DOI:10.1109/VETECE.2002.1040495, 1647–1651, 2002.
- [2] Burleigh, S.; Hooke, A.; Torgerson, L.; Fall, K.; Cerf, V.; Durst, B.; Scott, K.; Weiss H. (2003); Delay-tolerant networking: an approach to interplanetary internet, *IEEE Commun. Mag.*, DOI:10.1109/MCOM.2003.1204759, 128–136, 2003.
- [3] Buttyan, L.; Schaoer, P. (2010); PANEL: Position-Based Aggregator Node Election in Wireless Sensor Networks, *Int.J. Distrib. Sens. Netw*, DOI:10.1.1.158.9102, 1–16, 2010.
- [4] Chan, H.; Perrig, A. (2004); An Emergent Algorithm for Highly Uniform Cluster Formation, *Wireless Sensor Network*, DOI:10.1007/978-3-540-24606-0\_11, 154–171, 2004.
- [5] Chiang, CC.; Wu, HK.; Liu, A.; Gerla M. (1998); Routing in clustered multi-hop mobile wireless networks with fading channel, *in Proc. of IEEE SICON 98*, 197–211, 1998.
- [6] Crepaldi, R.; Bakht, M.; Kravets R. (2012); QuickSilver: application-driven inter- and intra-cluster communication in Vanets, *in Proc. ACM MobiOpp 02*, DOI:10.1145/2159576.2159591, 69–76, 2012.
- [7] Dang, H.; Wu H. (2010); Clustering and Cluster-based Routing Protocol for Delay-tolerant Mobile Networks, *IEEE Trans. Wireless. Comm*, DOI:10.1109/TWC.2010.06.081216, 9(6), 1874–1881, 2010.
- [8] Ding, P.; Holliday, J.; Celik A. (2005); Distributed Energy Efficient Hierarchical Clustering for Wireless Sensor Networks. *in Proc. IEEE DCOSS*, DOI:10.1007/11502593\_25, 322–339, 2005.
- [9] Fall, K. (2003); A delay-tolerant network architecture for challenged internets, *in Proc. of The 2003 conference on Applications, technologies, architectures, and protocols for computer communications (SIGCOMM '03)*, DOI:10.1145/863955.863960, 27–34, 2003.
- [10] Fan, CS. (2016); HIGH: A Hexagon-based Intelligent Grouping Approach in Wireless Sensor Networks, *Advances in Electrical and Computer Engineering*, DOI:10.4316/AECE.2016.01006, 16(1), 41–46, 2016.
- [11] Fang, W.; Li, S.; Liang, X.; Li, Z. (2012); Cluster-based Data Gathering in Long-Strip Wireless Sensor Networks, *Advances in Electrical and Computer Engineering*, DOI:10.4316/AECE.2012.01001, 12(1), 3–8, 2012.

- 
- [12] Golanski, M.; Schoeneich, RO.; Siwko, M. (2010); The algorithm for distribution of large-size data in the Wireless Ad-hoc Sensor Network, *in proc. Concepts and Implementation for Innovative Military Communications and Information Technologies, Military University of Technology*, 577–584, 2010.
- [13] Heinzelman, WR.; Chandrakasan, A.; Balakrishnan, H. (2000); Energy-Efficient Communication Protocol for Wireless Microsensor Networks. *in Proc. HICSS 07*, DOI:10.1109/HICSS.2000.926982, 10–19, 2000.
- [14] Iwata, A.; Chiang, C-C.; Gerla, M.; Chen, TW. (1999); Scalable Routing Strategies for Ad hoc Wireless Networks, *IEEE Journal on Selected Areas in Communications*, DOI:10.1109/49.779920, 17, 1369–1379, 1999.
- [15] Jacquet, P.; Muhlethaler, P.; Clausen, T.; Laouiti, A.; Qayyum, A.; Viennot, L. (2001); Optimized Link State Routing Protocol for Ad Hoc Networks, *IEEE Trans. Wireless. Comm.*, DOI:10.1109/INMIC.2001.995315, 62–68, 2001.
- [16] Jung, S.; Han, Y.; Chung, T. (2007); The Concentric Clustering Scheme for Efficient Energy Consumption in the PEGASIS. *in Proc. IEEE ICACT 07*, DOI:10.1109/ICACT.2007.358351, 260–265, 2007.
- [17] Kim, SK.; Yoon, JH.; Lee, J.; Yang, SB. (2015); HCS: hierarchical cluster-based forwarding scheme for mobile social networks, *Wirel. Netw.*, DOI:10.1007/s11276-014-0876-x, 21(5), 1699–1711, 2015.
- [18] Koutsonikola, D.; Das, S.; Charlie, HY.; Stojmenovic, I. (2010); Hierarchical geographic multicast routing for wireless sensor networks, *Wirel. Netw.*, DOI:10.1007/s11276-008-0146-x, 16(2), 449–466, 2010.
- [19] Li, C.; Ye, M.; Chen, G.; Wu, J. (2005); An Energy Efficient Unequal Clustering Mechanism for Wireless Sensor Networks, *in Proc. IEEE MASS 05*, DOI:10.1109/MAHSS.2005.1542849, 596–604, 2005.
- [20] Lindgren, A.; Doria, A.; Schelen, O. (2003); Probabilistic routing in intermittently connected networks, *ACM SIGMOBILE Mobile Computing and Communications Review*, DOI:10.1145/961268.961272, 19–20, 2003.
- [21] Lindsey, S.; Raghavendra, C.; Sivalingam, LM. (2002); Data gathering algorithms in sensor networks using energy metrics, *IEEE Trans. Parallel Distrib. Syst.*, DOI:10.1109/TPDS.2002.1036066, 13(9), 924–935, 2002.
- [22] Liu, X. (2012); A Survey on Clustering Routing Protocols in Wireless Sensor Networks. *Sensors*, DOI:10.3390/s120811113, 12(8), 11113–11153, 2012.
- [23] Luo, H.; Ye, F.; Cheng, J.; Lu, S.; Zhang, L. (2005); TTDD: Two-tier data dissemination in large-scale wireless sensor networks, *Wirel. Netw.*, DOI:10.1007/s11276-004-4753-x, 11(1), 160–175, 2005.
- [24] Loscri, V.; Morabito, G.; Marano, S. (2005); A Two-Level Hierarchy for Low-Energy Adaptive Clustering Hierarchy, *in Proc. IEEE VTC -5*, DOI:10.1109/VETECF.2005.1558418, 1809–1813, 2005.
- [25] Manjeshwar, E.; Agrawal, DP. (2001); TEEN: A Routing Protocol for Enhanced Efficiency in Wireless Sensor Networks, *in Proc. IEEE IPDPS 01*, DOI:10.1109/IPDPS.2001.925197, 2009–2015, 2001.

- 
- [26] Manjeshwar, E.; Agrawal, DP. (2002); APTEEN: A Hybrid Protocol for Efficient Routing and Comprehensive Information Retrieval in Wireless Sensor Networks, *in Proc. IEEE IPDPS 02*, DOI:10.1109/IPDPS.2002.1016600, 195–202, 2002.
- [27] Murugunathan, SD.; Ma, DCF.; Bhasin, RI.; Fapajuwo, AO. (2005); A Centralized Energy-Efficient Routing Protocol for Wireless Sensor Networks, *IEEE Radio Commun*, DOI:10.1109/MCOM.2005.1404592, 43(3), 8–13, 2005.
- [28] Schoeneich, RO.; Golanski, M. (2007); Mesh Cluster Based Routing Protocol: Enhancing Multi-hop Internet Access using Cluster paradigm, *in Proc. EUROCON, 2007. The International Conference on Computer as a Tool*, DOI:10.1109/EURCON.2007.4400318, 962–965, 2007.
- [29] Spyropoulos, T.; Psounis, K.; Raghavendra, CS. (2005); Spray and wait: an efficient routing scheme for intermittently connected mobile networks, *in Proc. of The 2005 ACM SIGCOMM workshop on Delay-tolerant networking*, DOI:10.1145/1080139.1080143, 252-259, 2005.
- [30] Soro, S.; Heinzelman, W. (2005); Prolonging the Lifetime of Wireless Sensor Networks via Unequal Clustering, *in Proc. IEEE WMAN 05*, DOI:10.1109/IPDPS.2005.365 ,236–243, 2005.
- [31] Su, YY.; Hwang, SF.; Dow, CR. (2008); An Efficient Cluster-Based Routing Algorithm in Ad Hoc Networks with Unidirectional Links, *Journal of Information Science And Engineering*, 24(5), 1409–1428, 2008.
- [32] Vahdat, A.; Becker, D. (2000); Epidemic routing for partially connected ad hoc networks. *Tech Report CS-2000-06*, DOI:10.1.1.34.6151, 2000.
- [33] Wei, C.; Yang, J.; Gao, Y.; Zhang, Z. (2011); Cluster-Based Routing Protocols in Wireless Sensor Networks: A Survey, *in Proc. IEEE ICCSNT 11*, DOI:10.1109/ICCSNT.2011.6182285, 1659–1663, 2011.
- [34] Yau, S.; Gao, W. (2007); Multi-hop clustering based on neighborhood benchmark in mobile ad-hoc networks, *Mob. Netw. Appl*, DOI:10.1007/s11036-008-0039-3, 12(5), 381–391, 2007.
- [35] Ye, M.; Li, C.; Chen, G.; Wu, J. (2005); EECS: An Energy Efficient Clustering Scheme in Wireless Sensor Networks, *in Proc. IEEE IPCCC.05*, DOI:10.1109/PCCC.2005.1460630, 535–540, 2005.
- [36] Younis, O.; Fahmy, S. (2004); HEED: A hybrid, energy-efficient, distributed clustering approach for ad-hoc sensor networks, *IEEE Trans. Mobile Comput.*, DOI:10.1109/TMC.2004.41, 3(4), 366–379, 2004.