Distributed Connected Dominating Set Election from Uniform Random to Power Law Network Graphs

Joseph P. Macker
Information Technology Division
Naval Research Laboratory

ABSTRACT
This work describes an analysis of applying different election heuristics to connected dominated set (CDS) algorithms within different topological graph taxonomies. Our initial motivation for this study was a general observation that a significant amount of research in mobile ad hoc network (manet) algorithm design and related election heuristic evaluation is based upon geometric random or other types of uniform random network models. Our hypothesis is that small world or scale-free network topology analysis can result in significantly different design feedback that may be suppressed by exclusive use of uniform random analysis. Previous research has shown that power law based networks more accurately represent types of real world and social networks and we believe it is important to understand the potential tradeoffs in distributed algorithm design for self-organizing networks beyond traditional random network models. We provide a concrete example by applying the existing specification for the essential-CDS algorithm used within the mobile ad hoc network (manet) Simplified Multicast Forwarding (SMF) and the OSPF Extensions for MANET MDR approach. We model random, small world, and power law based network types based upon well known models. We demonstrate that the network graph taxonomy used in modeling affects the design analysis such as understanding the resulting effectiveness of various distributed election heuristics. Future work is planned to extend these initial results to more dynamic analysis.

BACKGROUND AND MOTIVATION
At present, algorithms for efficient unicast and multicast wireless network routing are being developed and actively specified within various technical communities (e.g., IETF manet, IEEE 802.x). In particular to this work, distributed connecting dominating set (CDS) algorithms are being developed and specified within existing manet work either for the purpose of more efficient, dynamic routing control plane flooding or for constructing multicast forwarding backbones. In addition, basic analysis of these algorithms is often done using random unit disk graph (UDG) models, or other geometric-based uniform random graphs. Other analysis work is often done using randomly generated mobility models, such as random vector waypoint based approaches available within well known simulation and emulation systems. While randomized mobility and UDG analysis provides important performance input, basing design decisions solely on analyzing these forms of network models may misrepresent typical performance when applied in more real world network topologies (e.g., networks with heterogeneous nodes/environments and natural clustering).

PROBLEM FOCUS AND DEFINITION
This work focuses on examining a particular CDS algorithm and its related maintenance heuristics motivated by the previous discussion. Most previous analysis of these types of algorithms has taken place using moderately scaled networks (e.g., 100 nodes) using uniform random or normalized mobility models. While our analysis is initially based upon non-dynamic models, our hypothesis is that network models exhibiting clustering or small world type features will exhibit additional design feedback that may not show up in more uniform model studies. We also hypothesize that additional design optimizations can be made without sacrificing efficiency performance and adaptation in more randomized scenarios. If true, this leads to a more win-win situation in advanced network design relating to social network and other communication network applications. We plan to employ dynamic models exhibiting non-uniform clustering behavior in future work efforts to study this area further.

In order to apply our study to a real world design problem, we have chosen to examine a presently specified algorithm building block within the manet research and standard specification.
community. The Essential Connected Dominating Set (E-CDS) algorithm has been identified as a distributed CDS algorithm with known application in both MANET-OSPF unicast routing extensions [E-CDS] and within the manet working group (WG) Simplified Multicast Forwarding (SMF) [SMF] specification (see Appendix A of that document). E-CDS is a fundamentally a distributed algorithm using localized neighborhood knowledge for electing and maintaining a CDS within a dynamic network. It uses at minimum \( k=2 \), where \( k \) is network hop distance, neighbor topology knowledge that can be collected via direct \( k=1 \) communication exchange. To work effectively, E-CDS requires an election heuristic to be specified as part of a network wide relay set update process. Figure 1 shows an example of an E-CDS forwarding backbone formed within a 50 network graph. The square labels are the elected relay set nodes and the solid lines are existing links between the relay sets forming a connected communication network backbone. All nodes in the network are at most one hop away from a relay set node within the backbone.

![Figure 1: E-CDS backbone formed with a 50 node network graph](image)

In present work a number of election heuristics have been discussed and implemented with various E-CDS and other distributed CDS or backbone algorithms but to our knowledge little or no work has been done to examine specific design and performance tradeoffs using multiple network taxonomy models. Present candidate heuristics for E-CDS election and dynamic maintenance include the following two primary schemes although more may be added in the future:

1. Neighborhood id-based election
2. Neighborhood degree-based election, with id tie-breaking method
3. Tuple scheme (Priority, degree, id): this scheme is equivalent to (2) when priority is constant. Priority can be thought as an additional weight value that can be applied.

For this study, we concentrate on specific studies comparing heuristic styles 1 and 2 above as primary election methods. Present designs generally support the capability to support style 3 elections involving several tuple parameters when desired, but it has not been clear when such election approaches provide benefit.

Considering our introductory hypothesis we also asked the following question at the beginning of this study:

1. What performance trends emerge if we compare CDS election heuristics across random-based, small world, power law network models?
2. How would similar trends appear across alternative non-random network topology models (small world vs. power law)?
3. How would other varying other factors within the models (node population, density) affect the overall results and trends?

While there are cases where real wireless networks may be reasonably represented by uniform random geometric models, there are certainly cases in which this is not true. We anticipate this to be more prevalent as network nodes become more heterogeneous in nature and when nodes are operated in non-homogeneous environments. In many ways, these types of network models are also more representative of networks arising from human interaction, system organization, and those occurring within many natural world systems. So results or observations from this type of distributed algorithm analysis may have implications beyond communication network routing design and can potentially be applied to other disciplines such as social network-based system organization, distributed directory optimization, cooperative robotic positioning, other structural adaptation during distributed network operations.

**BACKGROUND OF GRAPH MODELS**

As mentioned, our initial study focuses on static graph analysis and formulates the modeling based
upon existing graph models representing various taxonomies. Historically related graph generation methods applied by theorists and recent trends are represented briefly in Figure 2. This is by no means exhaustive but represents a several major models used by researchers including relatively recent small world and power law models. We shall apply variants of these models in our studies along with the UDG model.

**Figure 2: Graph Generation Models**

**Uniform Random Network Graph Models**

In mathematics, a random graph is a graph that is generated by some random process and is often used to study and analyze network-based problems. Different uniform random graph models produce different probability distributions on graphs. Most commonly studied and used historically is the Erdos–Renyi model, denoted $G(n,p)$, where $n$ is the number of vertices and in which every possible edge occurs independently with probability $p$. A closely related uniform random model, denoted $G(n,M)$, assigns equal probability to all existing graphs with exactly $M$ edges and $n$ vertices.

Another graph model that is extensively applied to model distributed wireless and mobile ad hoc communication networks and analyze distributed protocols is the uniform random geometric model. A random geometric graph is an undirected graph constructed by placing $n$ vertices using an independent, uniform random process within some bounded surface or region, often a two-dimensional square or rectangle in related research. Then, two vertices, $x$ and $y$, are connected if and only if the distance between them is at most some range unit $r$, $d(x,y) =< r$. This is often referred to as the UDG model as well.

**Small World Graph Models**

The Watts and Strogatz model is a random graph generation model that produces graphs with small-world properties. Small world properties in a graph exhibit short average path lengths and higher clustering. It was proposed by Duncan J. Watts and Steven Strogatz in their joint 1998 Nature paper [WS98]. While there are variants of this generation algorithm the fundamental Newman-Watts-Strogatz approach we will apply works as follows: First create a ring using $n$ nodes. Then each node in the ring is connected with its $k$ nearest neighbors ($k-1$ neighbors if $k$ is odd). Then shortcuts are created by adding new edges as follows: for each edge $u-v$ in the underlying “$n$-ring with $k$ nearest neighbors” with probability $p$ add a new edge $u-w$ with randomly-chosen existing node $w$. In contrast with the initial Watts-Strogatz model, in the Newman-Watts-Strogatz variant no existing edges are removed when new edges are added [NX].

**Power Law or Scale Free Graph Models**

Interest in scale-free networks has also blossomed in the last decade and began around 1999 with work by Albert-László Barabási and colleagues at the University of Notre Dame. Examining a portion of the Internet-based Web topology, they discovered that some nodes, which they called “hubs”, had significantly more connections than others and that the network as a whole had a power-law distribution if examined as a function of the number of links connecting to any one node.

The Barabási-Albert preferential attachment model is a technique developed to generate power law graphs. In 2002, Holme and Kim demonstrated that the Barabási-Albert generation algorithm can suffer from fragmentation and other undesirable effects as certain growth condition occur using their preferential treatment method. They enhanced the model to include a probabilistic attachment component and produced approximate average clustering. We apply their model later in our study to examine power law type network topologies.

Figure 3 contains visual examples of models generating a 25 node network using a similar probabilistic coefficient=0.4 where applicable. The graphical layouts use stretch models so the geometric model represents the resultant graph but
do not depict the initial physical placement used in
generation as in the geometric case.

![Example Generated Graphs using Study Models](image)

**Figure 3: Example Generated Graphs using Study Models**

**USE OF NETWORKX TOOLKIT**
To carry out the analysis in this study we take
advantage of a well-established network theory
programming toolkit, networkx [HSS08].
Networkx provides a library of graph theory tools
including a number of the widely used and
published graph generation models used by
theorists. We use the following graph generation
modules from networkx in our studies [NX].
- erdos_renyi_graph
- random_geometric_graph
- newman_watts_strogatz_graph
- powerlaw_cluster_graph

**OUR E-CDS MODELING IN NETWORKX**
We developed python code to represent the
distributed E-CDS algorithm as specified within
the Appendix of the Simplified Multicast
Forwarding (SMF) draft specification version 09
[SMF]. We then extended the existing set of
networkx graph processing methods and
approaches to provide a generic CDS output
module. Given a graph as input the CDS code
module returns a resulting CDS relay set or
vertices elected. The code operates on non­
connected graphs as well and will return results
operating on all subgraphs. This properly
simulates the distributed nature of true E-CDS
operation in the real world. The code provides an
indicator that it has detected fragmentation so that
such experimental cases may be extruded or
flagged in analysis if desired. We also developed
post processing to check CDS set correctness by
testing the actual “connectedness” and “CDS
coverage” across the subgraph(s). We also provide
the ability to configure different election heuristics
within the CDS generation module.

**STUDIES AND ANALYSIS**
The following section presents the studies
conducted using the distributed CDS algorithm
code for generating relay set backbones given a set
of input graphs. We present several case studies
for 100 node scenarios. There are two
performance graphs presented in each case: a
graph for E-CDS election using an ID-based
heuristic and second graph using an alternative
{degree, ID} tuple election heuristic that considers
degree of a node. In the second algorithm case,
neighborhood nodes with equal degree use ID
tuple values to break degree-based election ties. In
networks where all degrees are equal the
algorithms are essentially equivalent. However,
as more varied clustering occurs within networks the
algorithms are expected to diverge. Throughout
our studies, the neighborhood dimension is \( k=2 \).
Each node is knowledgeable of directly connected
edges and those edges connecting 2-hop neighbors.
This minimal information simulates protocols that
operate within wireless mobile ad hoc network
collecting 2-hop neighborhood information using a
1-hop HELLO protocols as in MANET-OSPF
extensions [E-CDS], OLSR [OLSRv2], or NHDP
[NHDP] specifications. In running this study it is
important to note that nodes are not aware of
existing edges between 2-hop neighbors. Such
additional local knowledge would provide more
efficient election of a distributed CDS but is not
representative of \( k=2 \) distributed algorithms.

Our studies concentrate on examining the efficient
tradeoffs in the average number of relays elected
for each algorithm variant averaged across a set of
\( N=4 \) randomized trials for each test point. The
coefficient values affecting density are varied
across the range to collect 20 data point averages
representing varying densities. Many of the
generation algorithms provide a coefficient value
that influences the density and/or clustering in the
networks. We use the clustering coefficient
analyzer provided with networkx to plot the
clustering coefficient trend along with each plot as
our coefficients are varied. This graph provides a representative look at clustering trends within the network as we vary other generator coefficient parameters. For the Erdos-Renyi model, as in Figure 4, there is a clear linear relationship between the coefficient value and the resulting clustering coefficient results.

EXPERIMENTS AND DISCUSSION
This section provides a summary of basic experiments performed and discusses results.

Erdos-Renyi Uniform Random Graph Study
To study Erdos-Renyi type graphs across a spectrum of cases we generated random graphs using the following parameters:

```
erdos_renyi_graph(numnodes=100, coeff [0.0, 1.0], randomseed)
```

In figure 4, we observe the fact that for many regions of lower coefficient values we observe some differences in resulting relay set efficiency between ID and degree-based election with degree election resulting in an overall lower relay set. Overall, the differences are slight especially as the graphs begin to contain more edges. In higher coefficient regions the results are essentially equivalent (i.e., coeff > 0.4).

For the purposes of this study those results can be ignored. As the graph test cases begin to become connected graphs, above coeff ~ 0.2, we observe regions of interest. As in the Erdos-Renyi case, we find little difference between the effectiveness of the two algorithm heuristics. Observable differences are slight throughout these tests. This is an interesting observation because such random geometric model variants are frequently used in research papers to characterize algorithm performance of distributed wireless ad hoc algorithms.

**Figure 5: Random Geometric Results**

Newman-Watts-Strogatz (Small World) Graph Study
To study small world type networks we apply a Newman-Watts-Strogatz graph generation model that is available within networkx and again studied ECDS operating on these graphs across a set of varying conditions. We generated graphs using the following parameters:

```
newman_watts_strogatz_graph(numnodes=100, k_degree=6, coeff [0.0,1.0],randomseed)
```

We observe immediately that the ECDS election algorithm variants are now producing significantly different results. The trends match intuitive prediction because, as this network begins to contain more advantaged nodes with more interconnected vertices, local degree-based election has a higher probability of electing these nodes in a distributed fashion. The significant degree of separation is interesting to observe and how it remains relatively significant through the coefficient range studied. Modifying the generator algorithm degree parameter will result in differing
results but overall we have observed that the general trends hold when studying these types of graphs in regions of interest.

Figure 6: Small World Graph Study
Powerlaw Clustering Graph Study
To study powerlaw type graphs we use the following parameters:
```
powerlaw_cluster_graph(numnodes=100, k_degree=6, coeff[0.0,1.0],randomseed)
```
ECDS election algorithm variants are now producing even more significantly different results vs. the small world graph case. ID election efficiency additionally suffers in powerlaw clustering graphs because there is small set of vertices with a significant large number of edges forming highly clustered nodes. The probability of electing these nodes with an ID scheme is decreased. We see the trends for performance remain constant through this experiment. What this means is that distributed Degree election is finding the key clustered nodes and electing them as part of the relay set where on average the ID election scheme is not able to do so even as additional edges are added in this model. This matches reasonable well with intuition but the actual degree of difference and trends provide useful insight.

Figure 7: Power Law Clustering Study
CONCLUSION
We developed a working model of the E-CDS algorithm within a toolkit for performing network graph analysis. The E-CDS algorithm is presently being applied within both specified unicast and multicast manet protocols. We carried out a study of this algorithm in a variety of graph taxonomies to examine performance trends as election heuristics were modified. Our hypothesis being that analyzing a distributed networking algorithm within non-uniform random networks might provide useful performance insight and potentially result in critical design feedback. Clustered and heterogeneous graph models are often not historically considered in basic network studies and algorithm development. We clearly show that non-uniform network graph generation models, both small world and powerlaw clustered models, produced profound differences in resultant relay set efficiency vs. election heuristics across a range of test and density conditions. We also directly demonstrate that these resultant differences can be masked by analysis that includes a range of densities but uses only uniform random generation models. Our results provide a novel contribution by analyzing election heuristics for distributed CDS algorithms in multiple graph taxonomies. We also believe our work provides insight into a set of more general analysis considerations that are useful in future work and design phases of distributed algorithms.

FURTHER WORK
We realize our work here is an initial, limited analysis and that there are many additional graph
generation algorithms and approaches that could be examined. In addition, our analysis here was limited to static comparison. Future work will address the challenge of mapping these studies to group-based mobility models involving clustering or heterogeneous link capabilities that dynamically form advantaged nodes within a real operational network. We also feel that further study of “real world” wireless deployment environments could determine more accurate generative model taxonomies for network analysis to supplement uniform random graph analysis. Such enhancements include proper terrain and environment modeling along with heterogeneous node capabilities. It is likely that such network or overlay structures may be better represented by small world or powerlaw relationships. Further work in this area is ripe for exploration.

ACKNOWLEDGEMENTS

The author would like to thank various colleagues for valuable discussion and those that helped validate the CDS algorithm itself through independent coding and analysis. Independent coding efforts included both Justin Dean and Brian Adamson of the Naval Research Laboratory. Also, invaluable thanks to Tom Wambold, Drexel University, for his efficiency improvements to the analysis code used in conducting this study. Also, thanks to Drs. Cynthia Dion-Schwarz and Santanu Das for supporting research related to furthering a more applied scientific understanding of distributed networking design.

REFERENCES


