Proceedings of the 2007 IEEE
Workshop on Information Assurance
United States Military Academy, West Point, NY, 20-22 June 2007

Efficient Distribution of Trust Authority Functions in Tactical Networks

Steffen Reidt and Stephen D. Wolthusen

Abstract— **In this paper we describe an algorithm for the distribution of trust authority functions such as key generation and distribution in tactical mobile ad hoc networks. Such networks cannot rely on existing infrastructures and must operate under severe resource constraints. Moreover, network partitioning and node failure, including Byzantine failures must be compensated in tactical networks. We propose the combination of metrics on both network state and beliefs or trust in other nodes to form a composite metric for use in a clustering algorithm. The effectiveness and other characteristics of this improved clustering algorithm are then evaluated and analyzed in a simulation environment, demonstrating a significant improvement over the baseline clustering algorithm.**

I. Introduction

Security architectures often tacitly assume the availability of cryptographic services, which may not be the case for mobile ad hoc networks (MANETs). Trust authority (TA) services form the basis for many advanced services, and the bootstrapping and their continued availability represent a significant challenge from both efficiency and security perspectives, particularly in hostile environments such as tactical networks. Such networks are self-organizing, self-discovering, rapidly changing in topology and devoid of dedicated infrastructural elements, and must cope with both active adversaries and limited resources such as energy, bandwidth, and computational power. Services to be provided, regardless of whether for traditional public key infrastructures (PKI) or for identity-based public key cryptography (ID-PKC) include the creation, distribution, and revocation of keys, as well as layered services such as authentication and authorization.

Recent research has investigated the issue of establishing a PKI on a subset of nodes in the network [1], [2] based on the use of cluster algorithms for the determination of cluster heads. Simultaneously, numerous authors have focused on the propagation of trust and developed models for establishing trust in MANETs [3], [4], [5], [6], [7]. In this paper, we report on the efficiency gained by combining such cluster algorithms with selected additional metrics, including trust, battery capacity of participating nodes, and metrics pertaining to the underlying network, namely cost of routing, bandwidth requirements, and desirable per-hop signal strengths. This provides robust criteria for the distribution of data (e.g. key material) and computation across nodes in a dynamic MANET as required for a distributed trust authority. The remainder of the paper is structured as follows: Section II provides a brief overview of cluster algorithms and trust metrics, which are used as specimens in the subsequent description of the TA clustering. Section III then introduces our distribution model, which contains the modified cluster algorithm and descriptions of the component metrics used in the algorithm. The model's implementation and embedding for evaluation purposes in the NS-2 [8] simulator are then described in section III, while section IV provides a brief analysis and evaluation of the implemented model. Finally, section V discusses our ongoing and planned extensions to the model and algorithms for efficient and robust TA distribution in tactical MANET environments.

II. Related Work

Overlays and Clusters as a Structuring Mechanism for Information Collection and Dissemination

In recent years clusters have been widely utilised to determine subsets of mobile ad hoc networks under the objective of saving energy [9], enhancing routing protocols [10], finding efficient flooding [11], [12], and broadcasting mechanism [13], or to generally build low-cost backbones [14]. Clusters have also been applied in recent research on distributing trust authorities in ad hoc networks [1], [2].

Bechler [1] establishes a security architecture using clustering and (k, n)-threshold cryptography, but does not consider trustworthiness. In each cluster, exactly one distinguished node, the cluster head, is responsible for establishing and organizing the cluster. Clusters are formed as geographically needed: If nodes cannot find existing clusters, they create some themselves, with existing clusters being merged and split on demand. Secret shares are distributed among the cluster heads, and thus are constantly (but not necessarily securely) spread over the network. A further drawback in Bechlers work is the significant relevance of gateway nodes, which act as connectors between neighbored clusters. As Bechlers simulation results illustrate, 34.2% of the overhead traffic is produced by the gateway nodes, whereas the cluster heads only produce 47.5% of the overhead traffic, although they incur the management of the security shares. Lin-Jiun [2] also builds a cluster-based security architecture for MANETs and avoids the

Royal Holloway, University of London, UK
Gjøvik University College, Norway and Royal Holloway, University of London, UK

ISBN 0-0000-0000-0/$10.00

issue of trust establishment assuming that every node has already exchanged a public key and a session secret key with its direct neighbors. Since we assume wireless data transfer, there is no reason why these initially exchanged keys should be trustworthy, calling the underlying assumptions into question.

For the purposes of TA distribution, conventional clustering suffers from the fact that cluster formation is heavily influenced by the initial geometry of the network, typically resulting in a central node of the cluster becoming cluster head, rather than the most trustworthy one. In addition, the same effect also leads to undesirable bunching of TA nodes, which increases the risk of compromise of larger groups. This is partially addressed by probabilistic approaches such as the one proposed by Zongpeng [15]. Here, every node participates in a communication backbone with a certain probability dependent on the number of its neighbors. Although this approach creates an energy-efficient backbone, it does not consider the energy and depletion levels of the nodes. While overlays represent a general organizing principle for creating node subsets of interests, the cost of creating and maintaining such networks is non-negligible. This suggests the possible use of multipurpose overlays to balance the costs across several applications. However, for the purposes of this investigation we are not interested in developing new cluster algorithms but rather to develop a *cluster metric* which incorporates the TA requirements into existing algorithms. The metric is intended as an open collection of parameters which can be extended as required; for the purposes of the discussion here, we concentrate on the aspects of *trust metrics* and constraints imposed by the tactical MANET environment itself, namely limited battery capacity, and RF interface constraints. Evaluation of the efficacy of the cluster metric is achieved by using the algorithm reported in [16] for max-min d-cluster formation in wireless ad hoc networks. This algorithm results in each node either being a cluster head itself or being at most d hops away from a cluster head. As d is configurable and the selection criterion in the basic algorithm is formed by the node identification numbers, we substitute this value by a *cluster metric* as described in section III-C.6. The following briefly reviews related work on trust metrics, as this partial metric has been the most intensely studied.

Trust Metrics

This section reviews a selection of trust metrics which have been proposed in recent years; while some of these have not been explicitly proposed in the form of metrics, we have adapted them to provide consistent terminology. All models share the use of a digraph-based representation with different vertex and edge valuation interpretations.

One of the first trust metrics was proposed by Zimmermann [17] in 1995 has remained popular. Here, nodes are keys of a public key system and the edges represent certificates. A user assigns a value from the set {`unknown, not trusted, marginally trusted, fully trusted`} to every key he retrieves. The reduction to only four different types of trust allows the model to be implemented easily. However, Maurer [18] showed that due to this simplicity the model may delivers counter-intuitive results in special scenarios. A seminal approach to define a trust metric in the form of a model for public-key certification, trust and recommendations was defined by Maurer [3] in 1996. Maurer established the syntax of `certificates, recommendation, trust` and `authenticity of public keys`, which form the axioms of his model. Based on these axioms, two intuitive inference rules are defined which permit drawing of transitive conclusions from a set of given axioms. Since this model is totally deterministic, Maurer inserts in a second step the consideration of confidence on a continuous scale between 0 and 1. This model is generic in the sense that it allows confidence values in a continuous scale from 0 to 1 and considers inferences of arbitrary depth and complexity. In order to enable a real implementation and a computation without exponential complexity, Caronni [5] suggested several possible simplifications. However, the model can also be considered quite basic regarding the choice of axioms. The set of axioms in the original version does e.g. not contain a time parameter, which is necessary for key revocation. Marchesini [19] addressed this issue and extended Maurer's model by axioms for `properties, time` and `domain` and thus provided numerous additional abilities of the system, including key revocation. In 2006 Bicakci [20] also investigated the incorporation of certificate revocation in this system and Gligor [21] discussed the need of additional parameters such as multiple types of evidence, negative evidence, and false evidence when using Maurer's model in ad-hoc networks. Recent further work on trust metrics includes research by Sun [6], who proposes two axioms for trust models, namely that (1) concatenation propagation of trust does not increase trust and (2) multipath propagation of trust does not reduce trust. Sun proposes two trust models which handle trust as a value between -1 and 1. Both models can return counter-intuitive results, since the concatenation of two negative trust values can in both models result in a positive value. Although Sun seems to break the first axiom himself, this is not the case, since he only considers the axioms with absolute values. The second axiom of Sun is contradicted in several other trust models. Abdul-Rahman [22] and Xiong [23] calculate the trust value as the average of the values calculated from different paths. According to this, an additional positive but low evidence will reduce the resulting trust value and thus break the second axiom.

In the case of numerous chains between the nodes, the trust value can grow arbitrarily, what leaves the reader with the question about the significance of the trust values. Present models in which the combination of a distribution algorithm and a trust metric is proposed, leave the exact definition of the trust metric as a separate and thus unanswered issue [24]. In this paper we propose the first distribution algorithm, that is configurable by loadable and accurate defined metrics. Especially the trust metric, which is a modification of Maurer's metric [3], provides a subtle tradeoff between accuracy and feasibility.

III. Model

A. Definitions

Ad-hoc networks are commonly modeled as a graph $G = (V, E)$, where V is the set of vertices and E the set of edges. For the purpose of investigating the convergence behavior of our model in section IV, we propose the following extensions:

Definition 1 (Belief Set) Let V be the set of all nodes, then the relation $R(t) : V \times V \rightarrow [0, 1] \subset \mathbb{R}$ contains all *quality factors* of the networks at a certain time. R can be identified as a matrix $R = (r_{ij}) \in M(n \times n; [0, 1] \subset \mathbb{R})$ and is called the *Belief Set*.

Definition 2 (Quality factor) The *quality factor* $r_{ij} \in [0, 1]$ describes the belief of node i about node $j's$ qualification for being a TA[1] node.

Definition 3 (TA configuration) Let $R = (r_{ij}) \in M(n \times n; [0, 1] \subset \mathbb{R})$ be the *Belief Set*, then a matrix $S = (s_{ij}) \in M(n \times n; \{0, 1\})$ is called *TA configuration* if:

$$\sum_{0 \leq i \leq n} s_{ij} = 1, \quad 0 \leq j \leq n \qquad (1)$$

$$s_{ij} = 1 \quad \Rightarrow \quad r_{ij} > 0 \qquad (2)$$

Thus a subset of nodes is called TA configuration if every node is connected exactly to one TA node.

B. TA Distribution Mechanism

A cluster algorithm is used to determine the subset of TA nodes in the MANET. Without loss of generality and as noted in section II, we use a modification of Amis' [16] algorithm for the initial implementation and evaluation of the metrics used in distributing TA services. The underlying principle of deterministic cluster algorithms is to have each node exchange information with immediate neighbors and decide whether it is to be a TA node itself or whether to accept a peer node as a TA node. If a node A accepts another node B as a TA node, then node B will be the connector to the TA for node A. In the case of Amis' algorithm this information exchange procedure is performed d times, what yields a network where every node has a maximal distance of d hops to its TA-connector. Amis describes the basic idea of his algorithm as follows:

Initially, each node sets its *winner*[2] to be equal to its own node id. This is followed by the *floodmax* phase.

Definition 2 (*floodmax*) - Each node locally broadcasts its *winner* value to all of its 1-hop neighbors. After all neighboring nodes have been heard from, for a single round, the node chooses the largest value among its own *winner* value and the values received in the round as its new *winner*. This process continues for d rounds.

In our extension of Amir's algorithm the *winner* value is represented by a quality factor instead of the node identity. Moreover, the base algorithm's approach of choosing its d-hop cluster head based on the decisions of neighboring nodes in round $d-1$ must be augmented since different nodes might hold different views about a node's TA qualification. The base algorithm is therefore extended as follows:

TA Cluster Algorithm:

- Each node collects the information broadcast by neighboring nodes and retains this until it is refreshed or exceeds its predefined lifetime. Cluster information with a `hopsToGo` value greater than 1 are pushed on the stack `forwardInfo`, whereupon the respective `hopsToGo` value is decreased by 1.
- In certain (possibly node-specific) time periods each node determines all quality factors about his known d-hop neighbors, choosing the node with the highest quality factor as its cluster head. If the node itself holds this value, or if another node has chosen him as cluster head, the node will itself be a TA node. The node then broadcasts the newly determined TA status, its additional information such as the battery level and `forwardInfo` to its neighbors. Every entry of the `forwardInfo` stack contains a parameter `hopsToGo`, which is indicating the number of forwarding hops and initially set to the cluster depth.

These modifications in Amis' algorithm yield a reduction in message complexity and hence also energy consumption. Owing to the stack `forwardInfo` cluster packets not being forwarded directly, the respective information is simply added to the next own packet. Even though this strategy decelerates the information exchange, the decisions of the nodes will be build on more up-to-date information than in Amis' original algorithm, since a decision is not longer made in d steps. A further advantage of the TA cluster algorithm is the strategy of completely local decisions, which are only based on collected information rather than on decisions of other nodes in previous rounds. This property permits building the cluster with very limited additional message traffic. In a network with an active data exchange and an underlying routing protocol, the information required by the cluster algorithm could simply be added onto other packets sent over the network. The actual message complexity in tactical networks does, however, depend on a number of parameters such as mobility and reachability (e.g. caused by topographical constraints) and will be investigated further in future work.

[1] For consistency, cluster heads are labeled TA nodes.

[2] *winner* is a TA node in the context of this paper.

C. Metrics

The choice of the cluster heads in our algorithm is based on the quality factors. In definition 2 the quality factor was fixed as a value in the continuous interval from 0 to 1. A quality factor $r_{ij} = 0$ means that a node i has no evidence about a node j, while a value of 1 perfectly qualifies node j as a TA node. In this section we develop several partial metrics, which will be combined to the *cluster metric*. Each partial metric is mapped onto the continuum $[0, 1]$, assuming no constraints are violated. In the case of a constraint violation of one or more partial metrics, the cluster metric will itself yield 0 and thus disqualify a node as TA node. For the purposes of this paper, partial metrics are merged to a cluster metric using a linear combination. This itself requires a linear and continuous mapping of the partial metrics and weighting for relative importance. The metrics discussed in this section are not exhaustive; the use of additional partial metrics is therefore discussed in section V.

C.1 Trust metric

The trust metric is the core of the cluster metric, since it induces the cluster algorithm to determine a set of essentially trustworthy TA nodes. In this paper, a modification of Maurer's [3] model for a public key infrastructure is used; however, both different trust model and valuations can be used, provided that the constraints described in section III-C are satisfied. Maurer's model consists of two parts, a deterministic and a probabilistic one. The basic model is, however, not suitable for implementation owing to its computational complexity and must be adapted in its deterministic part as described in the following section:

- **Deterministic part** The deterministic part defines the parameters, which are considered by the model, and defines inference rules for these parameters. Maurer labels the parameters as *statements* which include the *Authenticity of public keys*, *Trust*, *Certificates* and *Recommendations.* Based on those statements Maurer defines two inference rules, which consider recommendations of arbitrary depth. If e.g. a node A believes in the authenticity of a node X and he also trusts X to administer certificates and X holds a certificate of Y, then A will also believe in the authenticity of node Y (see [3] for details).

The first simplification of [3] yielding a reduction of complexity especially in the computations of the probabilistic part, is to restrict the trustworthiness statements to level 1, disallowing the use of second-hand evidence. For the purpose of building a pure trust model, we will also redefine the *statements* in [3] as follows:

– *Trust.* $Trust_{X,Y}$ denotes X's belief that a particular entity Y is a member of a friendly party and thus trustworthy for forwarding information and being a TA member.

– *Distrust.* $Distrust_{X,Y}$ denotes X's belief that a particular entity X is generally *not* trustworthy for forwarding information or being a TA member.

– *Authenticity of public keys.* $Aut_{A,X}$ denotes Alice's belief that a particular public key P_X is authentic.

The statements $Trust$ and $Distrust$ are the central parameters in our model. For the purpose of regarding *negative evidence* in a deterministic model it is necessary to define an additional parameter for distrust. Further *statements* such as Aut, that might deliver information about a node's trustworthiness can also be defined. Limiting evidence forwarding to level 1 and the above statements, inference rules are defined as follows:

$$Trust_{A,X},\ Trust_{X,Y}\ \vdash\ Trust_{A,Y} \qquad (3)$$

$$Trust_{A,X},\ Distrust_{X,Y}\ \vdash\ Distrust_{A,Y} \qquad (4)$$

$$Trust_{A,X},\ Aut_{X,Y}\ \vdash\ Aut_{A,Y} \qquad (5)$$

Rules (3) and (4) represent the forwarding of trust information over one hop, while (5) shows the mechanism to include additional statements in the model. The statement *Distrust* and additional rule (4) are necessary, since in the deterministic model every *statement* can only have the value 0 or 1, whereas we wish to model the three levels of trust "indifferent", "trusted" and "not trusted".

- **Probabilistic part** The deterministic model part defined all parameters of the trust model as fixed statements and inference. This allowed the deduction of all implicitly available statements. The probabilistic part adds the notion of uncertainty to statements in a continuous certainty range $[0, 1]$ with events assumed to follow the Laplace hypothesis. Every event is true only with a certain probability, and the core of the probabilistic part is to determine the certainty of the inferred events (statements). The following provides a brief summary of the model, for details on the base model refer to [3].

The set of statements which are contained in a node's A view is denoted by $View_A$. The closure of $View_A$ under the inference rules (3)–(5) is then labeled with $\overline{View_A}$, and contains the whole statement knowledge of node A including inferred statements. Since every statement shall be certain in a range from 0 to 1, the certainty of a statement is represented by the probability that this statement is true and the probability $P(S \in \overline{View_A})$ is labeled *confidence value.* The probability of an inferred statement S from node A is the probability of this statement being inferable from statements included in $View_A$, i.e. that $S \in \overline{View_A}$. With $\mathcal{S}_\mathcal{A}$ denoting the power set of $View_A$, the confidence value $conf(S)$ for a statement S can be defined as $\mathrm{conf(S) = P(S} \in \overline{View_A}) = \sum_{\mathcal{V} \subseteq \mathcal{S}_\mathcal{A} : \mathcal{S} \in \overline{\mathcal{V}}} \mathrm{P}(\mathcal{V})$.

The model defined so far allows to specify arbitrary dependencies between the statements in $\mathcal{S}_\mathcal{A}$. Having limited the level of inferences to 1, $P(\mathcal{V})$ can be computed as:

$$P(\mathcal{V}) = \prod_{\mathcal{S} \in \mathcal{V}} p(S) \cdot \prod_{S \notin \mathcal{V}} (1 - p(S))$$

Finally, the probability $p(S)$ for a derived statement S can be obtained as

$$p(S) = \mathrm{conf(S)} = \sum_{\mathcal{V}\subseteq\mathcal{S}_\mathcal{A}:\mathcal{S}\in\overline{\mathcal{V}}} \prod_{\mathrm{S}\in\mathcal{V}} \mathrm{p(S)} \cdot \prod_{\mathrm{S}\notin\mathcal{V}} (1-\mathrm{p(S)})$$

where the most costly, but due the limitation to inference level 1 still practical, computable part is the determination of the set $\{\mathcal{V} \subseteq \mathcal{S}_\mathcal{A} : \mathcal{S} \in \overline{\mathcal{V}}\}$.

For the purposes of this paper, the trust metric is defined by the statements *Trust* and *Distrust*. Trust in another node, in the context of building a TA, means the trust in the node's ability to be a TA member. Without loss of generality we assume all nodes to be equal in computational power as a simplification, hence trust is a measure for the belief that a node belongs to a friendly party. A crucial point in every trust system is the initial determination of trust. Without any knowledge of the network in an ad hoc environment, nodes could build trust assumptions on positive experiences such as a faultless connections. Since this can be subverted easily, we assume that trust is anchored by physical contact during the scenario.

Determining one trust value as input for the cluster metric requires that both the *confidence values* for *Trust* and the *Distrust* are combined to one trust factor $1 \geq t_f \in \mathbb{R}$. Every strategy that overstates one of the values would provide a potential point of attack. In the case of a strong effect of the Distrust value for example, an attacker could spread negative evidence about a node's neighbors and thus isolate the node from all its friendly neighbors. In order to minimize the ability of such attacks, we calculate the final trust factor f_t as $f_t = \mathrm{conf(Trust)} - \mathrm{conf(Distrust)}$.

C.2 Signal strength metric

In order to avoid a permanent transmission breakdown between a node and its TA connection, its desirable to choose a nearby node as TA connection. Since the distance between two nodes does not necessarily represent their connection quality, we choose the signal strength as a measure for the nearness of nodes. The signal strength is commonly specified in dBm and the benchmark data are provided by the maximal transmission power (100mW = 20dBm using the IEEE 802.11 standard as an example) and the threshold for the minimal required receiving power of -80dBm [25]. Since dBm already provides a logarithmized and thus feasible measure for the original mW values, we use the dBm values to define the signal strength factor f_s for a signal strength $s[dBm]$ as $f_s = \frac{s+80}{100}$.

C.3 Energy metric

Limited battery power is one of the major constraints in mobile ad-hoc networks. Since TA nodes generally perform a higher interaction with their neighbors than ordinary nodes, its desirable to choose TA members with a sufficient battery level. Most modern battery systems provide a direct or indirect metric based on the voltage of the batteries decreasing with the percentage of discharge $disc \in [0,1] \subseteq \mathbb{R}$ proportional to $1 - \sqrt[3]{disc}$ [26], we define the energy metric as $f_e = 1 - \sqrt[3]{disc}$.

C.4 Routing metric

Although the set of TA nodes, which is determined by our cluster algorithm, would provide a suitable routing backbone, the model is also intended to fit into a network with a preselected routing protocol. As stated in section III-B, a TA overlay network might be bootstrapped without performing additional data transfer. Under the premise of an existing routing protocol we define the *routing metric* to take advantage of already established routes. For this purpose we use the number of destination nodes rdn (routing destination nodes), that a node has reached within a certain time period rtp (routing time period), as a measure for its activity in the routing process. The value for rtp needs to be defined depending on the routing protocol, as well as a value rpn (routing perfect node) for the number of reachable destination nodes, i.e. a node with a *routing factor* of 1. According to this conventions, we define the *routing value*, which is the output of the *routing metric* similarly to the energy metric as follows:

$$f_r = \begin{cases} 1 - \frac{rpn-rdn}{rpn} & , rdn \leq rpn \\ 1 & , rdn > rpn \end{cases}$$

However, if there is no predefined routing protocol and the routing is performed using the TA nodes as backbone, then this metric is not required and hence not included in the cluster metric.

C.5 Bandwidth metric

While the routing metric encourages the concentration of data transfer to a small number of nodes, this can exceed the nodes' bandwidth. In order to avoid delays or dropped packets, the *bandwidth metric* measures the load of a node regarding its available bandwidth. As before, we use the IEEE 802.11g standard for our example without loss of generality. In 802.11g the data rate at a point of time is dependent on the signal strength and varies between 8 values from 6 Mbps[3] to 54 Mbps. We label these values $\mathrm{dr_1}$ (data rate 1) to $\mathrm{dr_8}$, where $\mathrm{dr_1}$ represents the lowest rate of 6 Mbps and $\mathrm{dr_8}$ the highest rate, respectively. Moreover, we define $\mathrm{dr_c} = \lfloor \mathrm{dr} \rfloor$ as the highest $\mathrm{dr_i}$ that is lower than dr, and d_i, $1 \leq i \leq 8$ denotes the respective data rate. Assuming that an available data rate at least *two* level above the minimum required one for the real data rate dr is most feasible, we define the *bandwidth factor* f_b as follows:

$$f_b = \begin{cases} 0 & , \mathrm{dr_c} \geq \mathrm{dr_i} \\ \frac{dr - \mathrm{dr_c}}{2\cdot(\mathrm{dr_{c+1}} - \mathrm{dr_c})} & , \mathrm{dr_{i-1}} \leq \mathrm{dr_c} < \mathrm{dr_i} \\ 0.5 + \frac{dr - \mathrm{dr_c}}{2\cdot(\mathrm{dr_{c+1}} - \mathrm{dr_c})} & , \mathrm{dr_{i-2}} \leq \mathrm{dr_c} < \mathrm{dr_{i-1}} \\ 1 & , \mathrm{dr_c} < \mathrm{dr_{i-2}} \end{cases}$$

[3] 1 Mbps = 10^6 bit per second

C.6 Cluster metric

All component metrics were designed to firstly return a value in $[0,1]$, or $(0,-1]$ in case of a violated constraint and to provide a linear correlation between their return value and its relative importance. The cluster metric finally combines all part metrics w.l.o.g. in a linear combination and returns the quality factor as fixed in definition 2. Let $\mathbb{M} = \{t, s, e, r, b\}$ be the set of indices of all part metrics and f_t, f_s, f_e, f_r, f_b be the respective return values based on the information of a node i about a node j at a certain time. Then the quality factor r_{ij} is calculated as:

$$r_{ij} = \begin{cases} \sum_{i \in \mathbb{M}} \lambda_i \cdot f_i & (f_i \geq 0 \; \forall i \in \mathbb{M}) \\ 0 & \text{otherwise} \end{cases}$$

with:

$$\sum_{i \in \mathbb{M}} \lambda_i = 1, \quad \lambda_i \geq 0 \; \forall i \in \mathbb{M}$$

The exact choice of the λ values is discussed in the following section IV and the incorporation of additional part metrics is covered in section V.

IV. Evaluation and Analysis

A. Implementation

For evaluating the cluster algorithm and eliciting the loading of the part metrics in the cluster metric, we have implemented the proposed model in the network simulator NS-2. NS-2 provides a energy model and a very basic mobility model, which allows the configuration of a certain number of nodes moving randomly in a fixed area. We ran simulations containing 6 to 50 nodes and configured the nodes to be highly mobile with a speed between 1 and 20 metres per second, while the broadcasting time in the cluster algorithm was chosen to 1 second. This configuration can be understood as a worst case configuration, since in tactical networks nodes are likely to move in coalition and thus with almost same speed in several groups.

The partial metrics are linearly combined in the cluster metric, which returns the final quality value as the belief of one node about another's node TA-qualification. All side conditions are fulfilled, since the cluster metric returns 0 in case of a broken side condition, i.e. at least one partial metric returns a negative value. In order to evaluate the functionality of the trust metric, we ran a simulation were 2 of 50 nodes were configured as enemies, while the nodes were able to identify each others as enemy or friend if they moved closer than 10 metres. After 100 seconds in a $700m$ x $700m$ area, all friendly nodes had a trust value about the both enemy nodes of -0.8 or smaller and thus did not choose them as TA nodes at all. In a second simulation 2 nodes were only configured as enemies for a period of 30 seconds, such that only two other nodes had physical contact with these hostile nodes during this 30 seconds. Even after 15 minutes the other nodes were changing their opinions about the two temporarily hostile nodes, what shows the brisance of intrusion detection especially for nodes with Byzantine behavior.

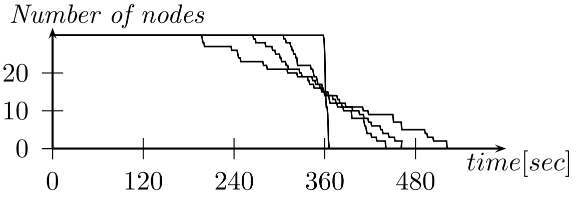

Fig. 1. Number of nodes with sufficient battery level.

Further simulations were performed to illustrate the quantitative effect of different configurations for the cluster metric. Figure 1 shows four different loadings between the node id and the battery level, while the other metrics were not considered, i.e. loaded as 0, and the initial battery level was lowly. The left graph presents the lifetime of the nodes for the loading (node id, battery level) = $(1, 0)$, which corresponds in the case of a 1-hop cluster to Amis original cluster algorithm. For this configuration the nodes start running out of energy after 200 seconds. The other three graphs show the lifetime of the nodes for the configurations $(2/3, 1/3)$, $(1/3, 2/3)$ and $(0, 1)$ and thus the influence of the energy metric. In the configuration (0,1), which is represented by the right graph, all nodes live as long as possible and run out of energy at almost the same time.

Due to permanent changes in the energy level and other parameters, which have an impact on the cluster metric, the quality value of the nodes and thus the TA members are swapping permanently. In future work we will investigate the incorporation of a metric, which measures the costs for selecting a new TA node to control this behavior. However, due to the simplicity of the mobility model, these results can only give a qualitative impression of the impact of the different metrics. More comprehensive mobility models and the evaluation of different loadings for the part metrics will be part of future work.

The rest of this chapter deals with the general quality of the cluster, especially with regards to the receipt of quality under less communication overhead.

B. Quality of Cluster Algorithm and TA Overlay

The quality factor forms the basis for the determination of the TA overlay. It is computed by the combination of several metrics as described in section III-C, and the constraints were considered by setting the quality to 0 if appropriate. Once the quality factors have been determined, the cluster algorithm can elect a near-optimal set of TA nodes while minimizing message traffic. Since every cluster algorithm is bound to local information exchange and to local decisions, there are two crucial constraints regarding the quality of the resulting cluster, namely

1. Limited information exchange in the cluster
2. Self-selection of cluster heads

In order to measure the effect of each of these two influences, we define the following three overlay configurations:
1. The *effectively chosen configuration* describes the set of TA nodes, which is selected by local decisions of the nodes and the information provided by the cluster algorithm. So as to approximately determine the respective subset of nodes, the cluster algorithm and the metrics have been implemented in the network simulator NS-2.
2. The *optimum cluster configuration* represents the best choice of the TA nodes based on information is provided by the cluster algorithm which is still valid. Since the determination of TA nodes is not influenced by local decisions but by the limited information provided by the cluster algorithm, the difference between the quality of this configuration and the effectively chosen configuration is a measure for drawback 2.
3. The choice of TA nodes in the *global optimum configuration* is based on the theoretically possible knowledge at a certain time, and thus differs from the *optimum cluster configuration* only in the information provided. The theoretically possible knowledge can thereby be determined by exchanging all information between those nodes which are in each others communication range. Configuration 2 will therefore converge to this configuration, if the time period of the information-broadcasting as part of the cluster algorithm converges to 0. Consequently, the difference between the quality of configuration 2 and 3 is a measure for constraint 1.

The definition of a *global optimum configuration* of the TA nodes as required for configuration 2 and 3 is ambiguous and might be dependent on the underlying information. In order to keep the battery power of all nodes on a reasonable level, the TA nodes could e.g. be determined under the premise of holding the minimum power of all TA nodes at the highest possible level, whereas in the case of trust the total amount of trust in the TA might be the essential value. However, we base our definition of the global optimum configuration of TA nodes on the *quality factor* (Definition 2) and propose the following definition:

Definition 4 (Quality of TA configuration) Let $Rows(S)$ be the number of non-zero rows of a matrix S, $\mathbf{1}$ the vector with 1-entries of length n and $*$ the multiplication element by element. Then the quality value $Qual_R$ of the TA configuration S for a Belief Set R is defined as:

$$Qual_R(S) = \frac{\mathbf{1}^{\mathrm{T}} \cdot (R * S) \cdot \mathbf{1}}{Rows(S)}$$

Let $\mathcal{S}$ be the set of all *TA configurations*, then the highest possible quality value $BestQual_R$ is defined as $BestQual_R = \max_{S \in \mathcal{S}} Qual_R(S)$. The *global optimum configuration BestConf*$_R$ for a Belief Set R is then one not necessarily unique TA configuration S with $Qual_R(S) = BestQual_R$

The crucial point in this definition is the quality value $Qual_R$, which is fixed as the average sum of quality values r_{ij} on the respective TA nodes. This definition tends to favor configurations with a small number of TA nodes, since a higher number of nodes generally decreases the mean. It can be seen from the simulations that the influence of local decisions (constraint 2) reduces the configuration quality $Qual_R$ on average by a factor of 1.2 to 1.4. The limited information exchange in the cluster causes a reduction of the quality, which is highly dependent on the mobility model and the time period of information broadcasting. Figure 2 illustrates the quality value $Qual$ and the related number of nodes for configurations 1 through 3. The nodes in the corresponding scenario were moving with a speed between 1 and $20m/s$, while the broadcasting time period of the nodes was set to $1s$. Although the time period of $1s$ is five times longer than proposed by Bechler [1] and Jiun [2], the quality loss caused by constraint 1 (difference between $Qual$ value of dotted and dashed line in figure 2) is marginal. The quality loss due to local decisions in the network (constraint 2) is caused by the larger number of TA nodes as can be seen from the distance between the solid and dashed lines in the lower graph of figure 2. The influence of constraint 2 appeared to decrease with the number of nodes. In a network of only 10 nodes, the quality loss due constraint 2 was on average 1.4, while in a scenario with 50 nodes this value decreased to 1.2.

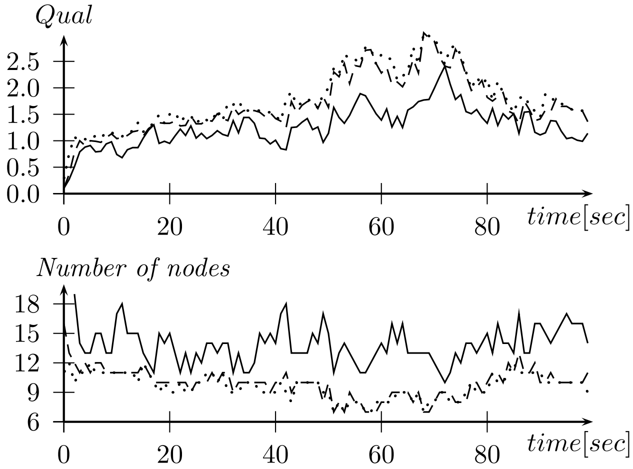

Fig. 2. Average quality and related number of TA members in a scenario of 50 randomly moving nodes in an area of 700m x 700m with a communication range of 100m and a cluster depth of 1. Configuration 1 = solid line, configuration 2 = dashed line, configuration 3 = dotted line.

V. Conclusions

In this paper we have described basic mechanisms for distributing TA services in tactical networks based on the development of combination metrics and their applications in a cluster-based algorithm which can be used both in the creation of service information overlays and together with additional lower-layer routing mechanisms. To permit the creation of TA services with limited assumptions while also taking advantage of already existing routing protocols, as typically provided in tactical networks, we investigated the determination of a feasible subset of TA nodes under the premise of limited communication overhead. Extension of an existing cluster algorithm with these metrics has demonstrated significant gains in efficiency while adding the ability to incorporate higher-layer properties such as trust; this area will be studied further in future work to ensure optimum selection of metrics and constraints, which can also be extended by including an element of dynamism.

Future work will investigate the configuration of the metrics for different scenarios and different mobility models as well as the characterization of higher-level TA services requiring distributed computation rather than mere secret sharing. For this purpose we will implement new mobility models, apply them to the already developed topographical model in NS-2 [27] and analyze the quality of the cluster in different scenarios regarding to a preferably minimal communication overhead. Furthermore, the integration of additional metrics for measuring the probability of compromise, topological position in the network and the costs for changing TA nodes will be evaluated. The following step in bootstrapping the security architecture is the creation of a (k, n)-threshold cryptography on the TA nodes. Our research in this area will include traditional public key infrastructures as well as identity based public key cryptography, and evaluate methods for distributed computation.

Acknowledgment

Research was sponsored by the U.S. Army Research Laboratory and the U.K. Ministry of Defence and was accomplished under Agreement Number W911NF-06-3-0001. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the U.S. Army Research Laboratory, the U.S. Government, the U.K. Ministry of Defence or the U.K. Government. The U.S. and U.K. Governments are authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation hereon.

References

[1] M.Bechler, H.J. Hof, D. Kraft, F. Pahlke and L. Wolf, “A Cluster-Based Security Architecture for Ad Hoc Networks,” in *INFOCOM 2004. Twenty-third AnnualJoint Conference of the IEEE Computer and Communications Societies*, (Hongkong), March 2004.

[2] Tsai Lin-Jiun, Lin Jen-Chiun and Lai Feipei, “SWARM: Secure Wireless Ad-hoc network Reliance Management,” in *Wireless Pervasive Computing, 2006*, pp. 1 – 6, 2006.

[3] Ueli Maurer, “Modelling a Public-Key Infrastructure,” in *Proc. 1996 European Symposium on Research in Computer Security (ESORICS’ 96)* (E. Bertino, ed.), vol. 1146 of *Lecture Notes in Computer Science*, pp. 325 – 350, Springer-Verlag, Sept. 1996.

[4] Michael K. Reiter and Stuart G. Stubblebine, “Authentication metric analysis and design,” *ACM Transactions on Information and System Security (TISSEC)*, vol. 2, pp. 138–158, May 1999.

[5] G. Caronni, “Walking the Web of trust,” in *IEEE International Workshops on Enabling Technologies: Infrastructure for Collaborative Enterprises (WET ICE’00)*, pp. 153 – 158, IEEE Computer Press,, 2000.

[6] Yan Lindsay Sun, Wei Yu, Zhu Han and K.J.Ray Liu, “Information theoretic framework of trust modeling and evaluation for ad hoc networks,” *IEEE Journal on Selected Areas in Communications*, vol. 24, pp. 305 – 317, Feb. 2006.

[7] G. Theodorakopoulos and J.S. Baras, “On Trust Models and Trust Evaluation Metrics for Ad Hoc Networks,” *IEEE Journal on Selected Areas in Communications*, vol. 24, pp. 318–328, Feb. 2006.

[8] Kevin Fall, Kannan Varadhan, *The ns Manual.* UC Berkeley, LBL, USC/ISI, and Xerox PARC, Oktober 2005.

[9] Carla-Fabiana Chiasserini, Imrich Chlamtac, Paolo Monti and Antonio Nucci, “An energy-efficient method for nodes assignment in cluster-based Ad Hoc networks,” in *Wireless Networks*, vol. 10, Springer, May 2004.

[10] D. Baker, A. Ephremides and J.A. Flynn, “The Design and Simulation of a Mobile Radio Network with Distributed Control,” *IEEE Journal on Selected Areas in Communications*, vol. 2, pp. 226 – 237, 1984.

[11] Taek Jin Kwon and Mario Gerla, “Efficient flooding with Passive Clustering (PC) in ad hoc networks,” in *ACM SIGCOMM. Computer Communication Review*, vol. 32, ACM Press, January 2002.

[12] Stefan Pleisch, Mahesh Balakrishnan, Ken Birman and Robbert van Renesse, “Routing and forwarding: MISTRAL:: Efficient flooding in mobile ad-hoc networks,” in *Proceedings of the seventh ACM international symposium on Mobile ad hoc networking and computing MobiHoc ’06*, ACM Press, May 2006.

[13] Foroohar Foroozan and Kemal Tepe, “A high performance cluster-based broadcasting algorithm for wireless ad hoc networks based on a novel gateway selection approach,” in *Proceedings of the 2nd ACM international workshop on Performance evaluation of wireless ad hoc, sensor, and ubiquitous networks PE-WASUN ’05*, ACM Press, October 2005.

[14] X.-Y. L. Yu Wang, WeiZhao Wang, “Distributed low-cost backbone formation for wireless ad hoc networks,” in *Proceedings of the 6th ACM international symposium on Mobile ad hoc networking and computing MobiHoc ’05*, ACM Press, May 2005.

[15] Li Zongpeng and Li Baochun, “Probabilistic power management for wireless ad hoc networks,” *Mobile Networks and Applications*, vol. 10, pp. 771–782, October 2005.

[16] Alan D. Amis, Ravi Prakash, Dung Huynh and Thai Vuong, “Max-Min D-Cluster Formation in Wireless Ad Hoc Networks,” in *Proceedings of the 21st Annual Joint Conference of the IEEE Computer and Communications Societies INFOCOM ’02*, pp. 32–41, 2000.

[17] P.R. Zimmermann, *The Official PGP User’s Guide.* Cambridge, 1995.

[18] Reto Kohlas and Ueli Maurer, “Confidence Valuation in a Public-key Infrastructure Based on Uncertain Evidence,” in *Proceedings of Public Key Cryptography 2000*, vol. 1751 of *Lecture Notes in Computer Science*, pp. 93–112, Jan. 2000.

[19] John Marchesini and Sean W. Smith, “Modeling Public Key Infrastructures in the Real World,” in *EuroPKI 2005. Second European PKI Workshop*, Springer-Verlag, June 2005.

[20] Kemal Bicakci and Bruno Crispo and Andrew S. Tanenbaum, “How to incorporate revocation status information into the trust metrics for public-key certification,” in *Proceedings of the 2005 ACM symposium on Applied computing SAC ’05*, ACM Press, March 2005.

[21] V. Gligor, “Trust Establishment, Trust Evidence, and Evaluation Metrics for Dynamic Taskforces,” tech. rep., University of Maryland, Nov. 2006.

[22] Alfarez Abdul-Rahman and Stephen Hailes, “A Distributed Trust Model,” in *Proceedings of the 1997 Workshop on New Security Paradigms (NSPW ’97)*, (New York, NY, USA), pp. 48–60, ACM Press, 1997.

[23] Li Xiong and Ling Liu, “Building Trust in Decentralized Peer-to-Peer Electronic Communities.”

[24] M. Virendra, M. Jadliwala M. Chandrasekaran and S. Upadhyaya, “Quantifying trust in mobile ad-hoc networks,” in *Integration of Knowledge Intensive Multi-Agent Systems*, pp. 65 – 70, IEEE Press, April 2005.

[25] Stefan Appel, “Lokalisierung von Knoten in mobilen Ad-hoc-Netzen ohne zusaetzliche Infrastruktur,” 2006.

[26] G. Danese and Leporati and R. Lombardi and M. Nucita and G. Pedrazzini and G. Ricotti, “An Instrument for the Characterization of Voltage and Temperature Profile in NiCd and NiMH Batteries,” in *23rd Euromicro Conference: New Frontiers of Information Technology – Short Contributions*, p. 0178, 1997.

[27] Steffen Reidt and Peter Ebinger and Stephen D. Wolthusen, “Resource-constrained signal propagation modeling for tactical networks.”.