Duplicate Address Detection in Wireless Ad Hoc Networks Using Wireless Nature

Yu Chen, Eric Fleury

INRIA

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Duplicate Address Detection in Wireless Ad Hoc Networks Using Wireless Nature

Yu Chen — Eric Fleury

N° 5841

Feb. 2006

Thème NUM

ISSN 0249-6399 ISRN INRIA/RR--5841--FR+ENG

Duplicate Address Detection in Wireless Ad Hoc Networks Using Wireless Nature

Yu Chen , Eric Fleury

Thème NUM — Systèmes numériques
Projets ARES/INRIA

Rapport de recherche n° 5841 — Feb. 2006 — 17 pages

Abstract: We consider duplicate address detection in wireless ad hoc networks under the assumption that addresses are unique in two hops neighborhood. Our approaches are based on the concepts of *physical neighborhood views*, the information of physically connected nodes, and *logical neighborhood views*, which are built on neighborhood information that is propagated in networks. Since neighborhood information is identified by addresses, inconsistency of these two views might be caused due to duplicate addresses. It is obvious that consistency of physical and logical views on each node's neighborhood is necessary for a network to have unique addresses, while the sufficiency depends on the types of information contained in views of neighborhood. We investigate different definitions of neighborhood views. Our results show that the traditional neighborhood information, neighboring addresses, is not sufficient for duplication detetion, while the wireless nature of ad hoc networks provides powerful neighborhood information in detecting duplication.

Key-words: duplicate address detection, wireless ad hoc networks, wireless nature

Unité de recherche INRIA Sophia Antipolis
2004, route des Lucioles, BP 93, 06902 Sophia Antipolis Cedex (France)
Téléphone : +33 4 92 38 77 77 — Télécopie : +33 4 92 38 77 65

Détection de duplication d'adresse dans les réseaux ad hoc en utilisant les caractéristiques sans fil.

Résumé : Pas de résumé

Mots-clés : Pas de motclef

1 Introduction

A wireless ad hoc network is a group of wireless nodes which cooperatively and spontaneously form a network. Each node has routing capabilities and communication is done in a multi-hop fashion. Such a network provides a flexible means of communication without using any existing infrastructure or centralized administration. Significant research in ad hoc networks has focused on efficient routing, the majority of which assume that nodes are configured *a priori* with a unique address before they communicate. Since in ad hoc networks nodes join and leave at will, automated address assignment is required to dynamically configure wireless nodes upon their entry into the network. In traditional networks, dynamic address assignment can be performed by a Dynamic Host Configuration Protocol (DHCP) [7] server. But this solution is not suitable in wireless ad hoc networks due to the unavailability of centralized servers. One alternative is to allow nodes to pick tentative addresses and then the uniqueness of picked addresses is checked by some duplicate address detection mechanism; if duplications are detected, nodes pick new tentative addresses.

In this work, we focus on *duplicate address detection* in wireless ad hoc networks. Works on duplication detection have been proposed previously (e.g., [3, 16, 17, 18]). Many approaches assume the existence of global unique identification. Under this assumption, duplication can be detected by propagating associations of identifications and addresses.However, there is no global identification which is truly unique; e.g., IEEE medium access control (MAC) addresses are not truly unique. One alternative is to create an identification randomly [3, 16]. The argument is that the probability of collision is small if the range of identifications is large enough. But propagating large-ranged identifications will cause large packet overhead. Thus relying on unique identification is not desirable in ad hoc networks. In our work, we consider detecting duplicate address based on *local* uniqueness: we assume addresses are *unique in two hops neighborhood*. This assumption is made due to two facts. First, since symmetry can prevent a problem to be solved in anonymous networks [2, 8, 9], some form of uniqueness is necessary; compared to the assumption of *global* unique identifications, our assumption is much weaker. Second, many algorithms have been proposed to assign addresses that are unique in two hops neighborhood (e.g., [10]).

We observe that protocols that are not aware of duplicate addresses behave as if all the packets from the same address are from the same node. For example, link state routing running on a node with address *ip* regards all the nodes that are connected to a node with address *ip* as its neighbors. Thus if duplicate address exists, the view of link state routing on the neighborhood is different from the physical neighborhood view. Based on this observation, we propose the concepts of *physical neighborhood views* and *logical neighborhood views*. Informally, a *physical neighborhood view* of a node is information of nodes physically connected to it; examples include the number of neighbors, addresses of neighbors and distances to each neighbor. A *logical neighborhood view* is built based on neighborhood information identified by *addresses*: a node with address *ip* considers all the nodes that connect to a node that has address *ip* as its neighbors and the view is built based on neighborhood information of all such “neighbors". For example, given a node that has address *ip*, the number of its neighbors in its physical view is the number of nodes *physically*

connected to it, and in its logical view it is the number of nodes *connected to a node that has address ip*. More detailed example will be given later in Figure 1.

We consider duplicate address detection by comparing the physical and logical neighborhood views of each node. Logical neighborhood views can be built if each node propagates to all the others the state of each of its links, identified by the addresses of its two ends. Since neighborhood information is required by most existing protocols and it usually contains two ends' addresses of each link, the overhead of our approaches depends on other information defined in neighborhood views. It is obvious that consistency of physical and logical views on each node's neighborhood is necessary for a network to have unique addresses, but whether it is sufficient depends on the types of information contained in neighborhood views. For example, if a neighborhood view is defined as the number of neighbors, it is sufficient only in a small class of networks.

In our work, we investigate different definitions of neighborhood views. We start from a simple definition of neighborhood views, which consists of neighboring addresses. The idea of detecting duplication by comparing neighboring addresses is proposed in PDAD-NH [18, 17]. However, without further investigation on the correctness, the authors claimed "in case the sender of the link state packet is a common neighbor of the nodes with the same address, the conflict cannot be detected by PDAD-NH. Thus, conflicts in the two hops neighborhood must again be detected by other means" [17]. In fact, we prove this approach fails in certain class of networks, even under the assumption of unique address in two hops neighborhood. We show this class of networks have the following properties: each existing address is assigned to the same number of nodes and there is a circle that has special properties. This class of networks might not be common in practice, but should not, therefore, be overlooked, since its existence indicates an important difference between wired and wireless networks. The properties of this class of networks provide strong hints for our second definition of neighborhood views, which also includes distance in x and y direction to each neighbor. We show that, under the assumption of unique addresses in two hops neighborhood, duplication can be detected if distance information satisfies certain accuracy, which means distance information can be represented in a small number of bits and overhead can be small. Note we do not assume the availability of strong position information such as GPS. Relative distance between neighboring nodes can be estimated by the signal strength or microwave [13, 19, 1]. Neighbor or stronger distance information is used in many works on wireless networks [4, 13, 14].

An interesting implication of our results is that the wireless nature of ad hoc networks provides powerful neighborhood information in breaking symmetry. In wired networks, typical neighborhood information is neighboring addresses, which is, as shown in our work, not sufficient to detect duplications. However, our results show that duplication can be detected with neighbor distance information. Note this information is available due to the wireless nature of ad hoc networks; it is not available in traditional wired networks.

2 Related Work

Dynamic Host Configuration Protocol (DHCP) [7] is commonly used for dynamic address assignment in traditional networks. Works on dynamic address assignment for ad hoc network include [11, 15, 12]. Solutions for duplication detection in ad hoc networks has been proposed previously (e.g. [3, 16, 17, 18]). In [3], each node has an fixed-length identifier which is randomly generated. A special message that includes nodes' address and identifier is diffused to the entire network; a node detects a duplicate address when it receives a message that has the same address as its own, but with a different identifier. Global unique or randomly generated keys are assumed in [16], in which duplication is detected by attaching key information in link state packets. The approach proposed in [16] successfully prevents packets from being delivered to wrong destinations. Most approaches for duplicate address detection require propagation of key information, which causes high packet overhead. Since lower protocol overhead is one of the most important design goals for wireless ad hoc networks, works have been done in achieving efficiency in terms of protocol overhead. Protocols proposed in [18] and [17] generate almost no protocol overhead: it detects address conflicts in a passive manner based on anomalies in routing protocol traffic. In particular, the idea of detecting duplication by comparing neighborhood information is proposed in approach PDAD-NH [18] [17]. However, no correctness proof is presented. In our work, we show this approach works in most networks, except a special class of networks; the existence of this class of networks indicates the different ability of wired and wireless networks in duplication detection using neighborhood information.

Much work has been done on anonymous networks in which no identifications are available [2, 8, 9, 6, 5]. Less work considers networks, especially wireless networks, with partial identifications. However, partial identification information, such as MAC addresses, are commonly available. Here we consider duplication detection using neighborhood information under the assumption of local uniqueness, which is not solvable in typical wired networks, but can be solved in ad hoc networks by using information provided by wireless nature.

3 System Model and Overview

We focus on stand-alone wireless ad hoc networks in which wireless nodes do not have access to a centralized server that could assign network-wide unique addresses. Instead of assuming global unique identifications, we consider duplication detection under the assumption that addresses are unique in two hops neighborhood. In our work, duplicate address is detected by each node comparing its *physical neighborhood view* and *logical neighborhood view*. In section 1, we have given an informal description of physical and logical neighborhood views. In the sequel, we focus on whether the consistency of physical neighborhood view and logical physical neighborhood view on every node is sufficient for a network to have unique addresses. If it is sufficient, at least one node will detect duplicate address and it can inform other nodes about it. We consider two definitions; each definition has its own assumptions on neighborhood knowledge.

Physical neighborhood views are built based on neighborhood knowledge that are assumed to be available, thus no packet overhead is caused. But building logical neighborhood views requires propagation of neighborhood information, which causes packet overhead. We assume each node that has address ip generates packets $\langle ip, ip', link\_state\rangle$ for each neighbor that has address ip'; the field $link\_state$ will be specified by each approach. We borrow the name from link state routing and call these packets as *link state packets*. Since neighborhood information is required by most protocols and how to propagate this information is out of the scope of this paper, we assume each node receives link state packets from all the other nodes without going into details of how these packets are propagated. Since most neighborhood information contains two ends' addresses of each link, we evaluate the overhead of each approach based on the packet complexity of field $link\_state$ in link state packets.

In the first approach, we assume neighboring addresses are available and neighborhood view is defined as a set of neighboring addresses. No overhead is introduced. We prove this information is not sufficient and this approach fails in certain class of networks; in this class of networks, each existing address is assigned to the same number of nodes and there is a circle in which the sequence of nodes' addresses consists of repeated patterns. Based on this property, we propose our second definition. We observe that, due to its wireless nature, neighbor distance information is available in ad hoc networks. In our second approach, neighborhood view is defined as distances in x and y direction to each neighbor, together with ends' addresses of each link. Overhead of this approach is distant information in link state packets. We show that duplication can be detected if nodes that have the same address are not too "close"; the meaning of "being close" depends on the accuracy of neighbor distance. The allowance of inaccuracy implies a small number of bits can be used to represent distance information. Based on the assumption that addresses are unique in two hops neighborhood, we show that only small overhead is required.

In the sequel, we first introduce in section 4 the concept of addresses map, which is used in the analysis of the two approaches. We show in section 5 that duplication cannot be detected based on neighboring addresses, and in section 6, we show it can be solved by using neighbor distance information, which is available in ad hoc networks due to the wireless nature.

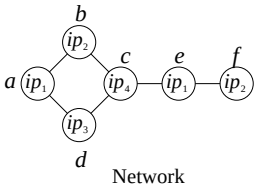
4 Addresses Map

In this section, we introduce the concept of *addresses map*. Informally, addresses map is a view of a network in which all the nodes with the same address are combined into one. An example is shown in Figure 1.

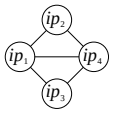
Definition 1 ***(Addresses Map)*** *Given a network G, its addresses map is a graph defined as follows:*

- *each vertex is a distinct existing address; and*

- *there is an edge between two addresses ip_1 and ip_2 iff $\exists$ link state packet $\langle ip_1, ip_2\rangle$ or $\langle ip_2, ip_1\rangle$.*



Network



Map of Addresses

node	links identified by end addresses
a	$\langle ip_1, ip_2\rangle$, $\langle ip_1, ip_3\rangle$
b	$\langle ip_2, ip_1\rangle$, $\langle ip_2, ip_4\rangle$
c	$\langle ip_4, ip_1\rangle$, $\langle ip_4, ip_2\rangle$, $\langle ip_4, ip_3\rangle$
d	$\langle ip_3, ip_1\rangle$, $\langle ip_1, ip_4\rangle$
e	$\langle ip_1, ip_4\rangle$, $\langle ip_1, ip_2\rangle$
f	$\langle ip_2, ip_1\rangle$

address ip	neighbors of ip in the map
ip_1	ip_2, ip_3, ip_4
ip_2	ip_1, ip_4
ip_3	ip_1, ip_4
ip_4	ip_1, ip_2, ip_3

Figure 1: An Example of Addresses Map

In the sequel, we use terms “address" and “edge" to refer to a vertex and a link in the addresses map respectively, and “node" and “link" to refer to a vertex and a link in a network respectively. We say a link connects two addresses ip and ip' if its two ends have addresses ip and ip'. The lemma below shows a *necessary* condition for a network to have duplicate address.

Lemma 1 *Given a network G in which addresses are unique in two hops neighborhood, if no circle exists in its addresses map, then no duplicate address exists in G.*

Proof. We assume in contradiction that duplicate address exists in G. Let nodes x and x' be the closest pair of nodes with the same address. Denoting the shortest path between x and x' by $\langle n_0, n_1, \ldots, n_{k-1}, n_k\rangle$, where $n_0 = x$ and $n_k = x'$, we have the following two properties: (1) $k \geq 3$ by the assumption of unique addresses in two hops neighborhood; (2) nodes in $\{n_0, \ldots, n_{k-1}\}$ have distinct addresses, since otherwise x and x' are not the closest pair of nodes that have the same address. Denoting the address of n_i by ip_i, there is a path $\langle ip_0, ip_1, \ldots, ip_k\rangle$ in the addresses map, which is a circle since $k \geq 3$ and $ip_0 = ip_k$. Contradiction! ■

Note the existence of circles in the addresses map is not a *sufficient* condition for duplication to exist in a network.

node	physical view	logical view	Consist.	node	physical view	logical view	Consist.
a	ip_2, ip_3	ip_2, ip_3, ip_4	False	b	ip_1, ip_4	ip_1, ip_4	True
c	ip_1, ip_2, ip_3	ip_1, ip_2, ip_3	True	d	ip_1, ip_4	ip_1, ip_4	True
e	ip_2,ip_4	ip_2, ip_3, ip_4	False	f	ip_1	ip_1, ip_4	False

Table 1: An Example: Views of Neighborhood of Network in Figure 1

5 Duplication Detection Using Neighboring Addresses

In this section, the only assumption on neighborhood knowledge is that each node knows its neighboring addresses. The view of a node's neighborhood is defined as *the set of neighboring addresses*. Since no information except ends' addresses of each link is required to build logical neighborhood views, link state packets have form of $\langle ip, ip' \rangle$ and no overhead is caused.

Definition 2 *Given a network and a node n in this network that has address ip, we define*

- *physical neighborhood view of $n \equiv$ the set of addresses of nodes that are physically connected to n.*
- *logical neighborhood view of $n \equiv \{ip' | \exists$ link state packet $\langle ip, ip' \rangle\}$*

The term "view" is used in this section according to this definition. Table 1 describes physical and logical neighborhood views of the network shown in Figure 1, in which this approach works. However, special symmetry can prevent this approach from detecting duplications. Counterexamples are given in Figure 2: all the nodes in Network 1 and Network 2 have consistent views, but duplications exist in both networks. Note these two networks have the same addresses map.

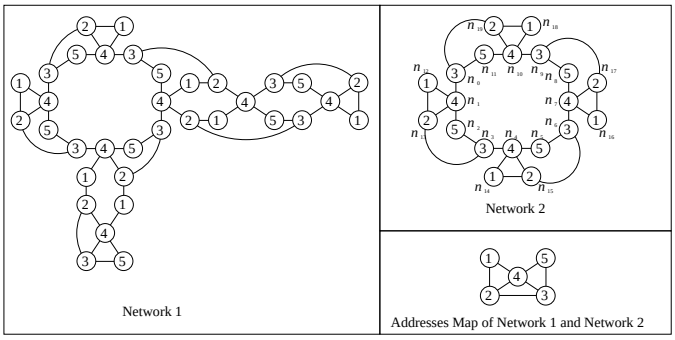


Figure 2: Examples of Networks in which views are consistent on every node

Now we investigate the properties of networks in which this approach fails. The lemma below shows that in such a network, addresses are distinct in the shortest path connecting any two different addresses.

Lemma 2 *Consider a network in which addresses are unique in two hops neighborhood and views are consistent on every node. Given any two addresses ip_x and ip_y, nodes in a shortest path that connects ip_x and ip_y have distinct addresses.*

Proof. We prove it by contradiction. Let $path = \langle n_0, n_1, \ldots, n_k \rangle$ be a shortest path that connects ip_x and ip_y. We denote the address of n_i by ip_i. Note $ip_0 = ip_x$ and $ip_k = ip_y$. Let i be the smallest index such that more than one node has address ip_i. Let i' be the second smallest index such that $ip_{i'} = ip_i$. We have $i' - i > 2$ by the assumption of unique addresses in two hops neighborhood. Note $i > 0$ since otherwise we have $ip_{i'} = ip_i = ip_x$ and path $\langle n_{i'}, \ldots, n_k \rangle$ is a shorter path that connects ip_x and ip_y. Letting $path_0 = \langle n'_0 \rangle$, where $n'_0 = n_{i'}$, we construct path $path_j$ by induction as follows (Figure 3):

- For some $j \in [0, i-1]$, assume $path_j = \langle n'_0, \ldots, n'_j \rangle$ satisfies (1) $\forall l \in [0, j]$, the address of n'_l is ip_{i-l}, and (2) $\forall l \in [1, j], n'_l \notin path$. It is obvious that $path_0$ satisfies these two properties.

- Given $path_j$, we now construct $path_{j+1}$ that also satisfies the above properties. Since n'_j and n_{i-j} have the same address and views are consistent on every node, n'_j and n_{i-j} has the same set of neighboring addresses. Since n_{i-j} has a neighbor n_{i-j-1} that has address ip_{i-j-1}, n'_j also has a neighbor that has address ip_{i-j-1}, denoted by n'. We now show $n' \notin path$. Otherwise, there are two cases:

  - If $n' = n_{i-j-1}$, then n_{i-j-1} is connected to n'_j and n_{i-j}. Since both n'_j and n_{i-j} have address ip_{i-j}, it contradicts to the assumption of unique addresses in two hops neighborhood.
  - Otherwise there is a node other than n_{i-j-1} in $path$ with address ip_{i-j-1}, which means n_{i-j-1} is a node that has duplicated address in $path$, which contradicts to that n_i is the first such node.

  Since $n' \notin path$ and n' has address ip_{i-j-1}, we can construct $path_{j+1}$ by letting n' be n'_{j+1} and appending it to $path_j$.

By the above construction, we get path $path_i$ in which the address of n'_i is $ip_0 = ip_x$ and $n'_i \notin path$, which implies $n'_i \neq x$. By replacing $n_0, \ldots, n_{i'}$ in $path$ by $path_i$, we have a path between n'_i and y with length $|path| - i' + i$, which is shorter than $path$ since $i' - i > 2$. Since the address of n'_i is ip_x, it contradicts to that $path$ is a shortest path that connects ip_x and ip_y. ■

The lemma below states that in such networks, given a path in which nodes have distinct addresses, there are t distinct paths that have same sequence of addresses, where t depends on the number of nodes that have the same address.

Lemma 3 *Consider a network in which addresses are unique in two hops neighborhood and views are consistent on every node. Given a path $path_0$ in which all the nodes have distinct addresses. Let t be the number of nodes that have address ip_0, where ip_0 is the address of*

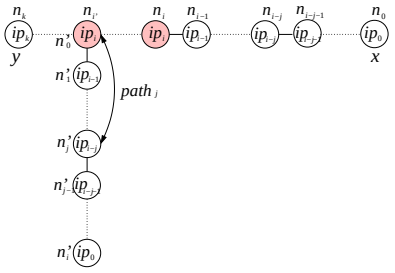


Figure 3: Proof of Lemma 2

the first node in $path_0$. Then there exist t distinct paths that have the same sequence of addresses as $path_0$.

Proof. Let $path_0$ be $\langle n_0^0, n_0^1, \ldots, n_0^k \rangle$. We denote the address of n_0^i by ip_i. We construct t paths by induction. For some $s \in [0, t-2]$, we assume there are $s+1$ paths, $path_0$, ..., $path_s$, that are distinct and have the same sequence of addresses as $path_0$. It is obvious it is true when $s = 0$. Now we construct $path_{s+1}$ as follows. We denote the nodes in $path_0$, ..., $path_s$ by $paths_{[0,s]}$.

Since there are t nodes that have address ip_0 and only $s+1 \leq t-1$ of them appear in $paths_{[0,s]}$, there is at least one node that has address ip_0 and is not in $paths_{[0,s]}$. Let this node be n_{s+1}^0.

Now we construct the rest of this path by induction. For $i \in [0, k-1]$, assume there is a path $\langle n_{s+1}^0, \ldots, n_{s+1}^i \rangle$ such that: (1) $\forall l \in [0, i]$, the address of n_{s+1}^l is ip_l; and (2) All the selected nodes, that is, $\{n_{s+1}^0, \ldots, n_{s+1}^i\} \bigcup paths_{[0,s]}$, are distinct. We select n_{s+1}^{i+1} as follows (Figure 4). Since views are consistent on every node and n_{s+1}^i and n_0^i have the same address, n_{s+1}^i and n_0^i have the same set of neighboring addresses. Since n_0^i has a neighbor n_0^{i+1} that has address ip_{i+1}, n_{s+1}^i also has a neighbor that has address ip_{i+1}, denoted by n'. Now we show n' is not among the selected nodes. Among all the selected nodes, only nodes in $\{n_0^{i+1}, \ldots, n_s^{i+1}\}$ have address ip_{i+1}. If n' is one of them, then n' is connected to n_l^i for some $l \in [0, s]$. So n' is connected to n_l^i and n_{s+1}^i, which have the same address ip_i. It contradicts to the assumption of unique addresses in two hops neighborhood. Since n' has not been selected and it has address ip_{i+1}, it can be selected as n_{s+1}^{i+1}.

■

Based on these two lemmas, the following theorem states that each existing address is assigned to the same number of nodes. An interesting implication is that a network with a prime number of nodes does not have duplicate address if views are consistent on every node. We define the *duplicate degree* of such a network as the number of nodes that take an existing address.

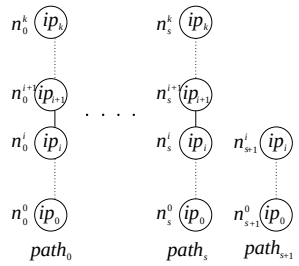


Figure 4: Proof of Lemma 3

Theorem 4 *Consider a network in which addresses are unique in two hops neighborhood and views are consistent on every node. For each address ip that exists in the network, the number of nodes that take ip as its address is the same.*

Proof. Assume in contradiction that there exist addresses ip_x and ip_y such that the number of nodes that have address ip_x is s and the number of nodes that have address ip_y is t, where $s, t \geq 1$ and $s > t$. Consider all the pairs x' and y' such that x' has address ip_x and y' has address ip_y. Let x and y be the closest pair among all these pairs. Let $path_0$ be the shortest path between x and y. By Lemma 2, the addresses of nodes in $path_0$ are distinct. By lemma 3, there are s paths with the same sequence of addresses as $path_0$ and all the nodes in these paths are distinct. So there are s nodes that have the same address as y, which contradicts to that only t, $s > t$, nodes has address ip_y. ■

The above theorem examines the connection between the *number* of nodes and that of addresses. Now we take a close look on the connection between the *topology* of a network and that of its addresses map. In particular, given a subgraph S_A of its addresses map, we examine the subgraph of a network that is "relevant" to S_A. Informally, a node is relevant if it has an address in S_A and a link is relevant if an edge connecting its two end addresses exists in S_A. We say such a subgraph is expanded by S_A. The formal definition is given below.

Definition 3 ***(Expanded Subgraph)*** *Given a network G and a subgraph S_A of its addresses map, we consider a subgraph S_G of G that satisfies:*

- *nodes in S_G are the nodes that have addresses in S_A, and*
- *there is an link between nodes x and y in S_G iff there is an edge between the address of x and the address of y in S_A.*

We say S_G is the subgraph that is expanded by S_A.

In the next theorem, we consider addresses that are organized in a circle in the addresses map. We show that the subnetwork that is expanded by it consists of a set of circles.

Furthermore, if duplicate address exists, there is a “minimal" circle in the addresses map whose expanded subgraph includes a circle that contains duplicate address; the existence of such a circle provides strong hints for our approach presented in the next section. Definition of “minimal circles" is given below. For example, in Network 2 of Figure 2, the circle $\langle 3, 4, 5, 3\rangle$ is minimal while the circle $\langle 1, 2, 3, 4, 1\rangle$ is non-minimal.

Definition 4 ***(Minimal Circle).*** *Given a graph G, a circle cir is minimal iff there exists a node x in cir such that cir is the shortest circle that contains x.*

Theorem 5 *Consider a network G in which addresses are unique in two hops neighborhood and views are consistent on every node. Given any circle $circ_{addr} = \langle ip_0, ip_1, \ldots, ip_k, ip_0 \rangle$ in the addresses map, the subgraph S_G of G that is expanded by $circ_{addr}$ consists of a set of circles, and each circle has the form of*

$$path_0 \circ path_1 \cdots \circ path_{s-1} \circ \langle n_0 \rangle$$

where $path_i$ is a path that has sequence of addresses $\langle ip_0, \ldots, ip_k \rangle$, n_0 is the first node in $path_0$ and $s \geq 1$ (Figure 5).

Furthermore, if duplicate address exists, then there exists a minimal circle in the addresses map whose expanded subgraph in G contains a circle that has $s > 1$ in the above form.

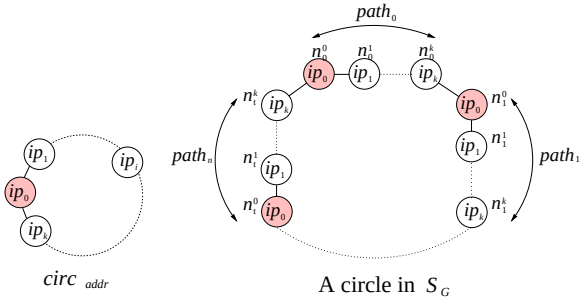


Figure 5: Theorem 5

Proof. Let t be the duplicate degree of G. We construct a subgraph S of G as follows:

- Note there is at least one path in G that has sequence of distinct addresses $\langle\ ip_0, \ldots, ip_k\ \rangle$. By Lemma 3, there are t distinct paths that has the same addresses sequence (Figure 6). By Theorem 4, each address is assigned to exactly t nodes, so these t paths contain all the nodes that have addresses in $circ_{addr}$.
- We denote by B (E resp.) the set of nodes at the beginning (end resp.) of these t paths and S is constructed by adding links between nodes in B and nodes in E to the t paths.

Note B (E resp.) is also the set of nodes that have address ip_0 (ip_k resp.). By Lemma 3, there are t distinct links in G that connect a node that has address ip_k and a node that has address ip_0. So each of these t paths is on a circle in S in the form defined in the theorem.

It is easy to see S contains all the node that have address in $circ_{addr}$ and a link between two end addresses of each link in S exist in $circ_{addr}$. For each edge in $circ_{addr}$, say $\langle ip, ip' \rangle$, t distinct links that connect ip and ip' are contained in S. Since exactly t nodes have address ip (ip' resp.) by Theorem 4, these t distinct links are all the links that connect ip and ip'. So $S = S_G$ by definition 3. Thus we prove the first part.

By Lemma 1, if duplicate address exists, there exists circle in the addresses map. We denote by $\langle x_0, \ldots, x_{k-1}, x_k \rangle$ the shortest path between the closest pair of nodes that have the same address. We have $k \geq 3$ by the assumption of unique addresses in two hops neighborhood. Note nodes in $\langle x_0, \ldots, x_{k-1} \rangle$ have distinct addresses, since otherwise x_0 and x_k are not the closest pair. Denoting the address of x_i by ip'_i, there is a circle in the address map, $\langle ip'_0, \ldots, ip'_{k-1}, ip'_k \rangle$, where $ip'_k = ip'_0$. Note this circle is minimal since otherwise there is a shorter path between x_0 and x_k. By Lemma 3, there are t distinct paths that have sequence of addresses $\langle ip'_0, \ldots, ip'_{k-1} \rangle$, including the path starting from x_0 and the path starting from x_k. We have proved all these paths are contained in a set of circles in the form defined in the theorem. Since x_{k-1} is connected to x_k, so the path starting from x_0 and the path starting from x_k are contained in the same circle. Thus the second part is proved. ■

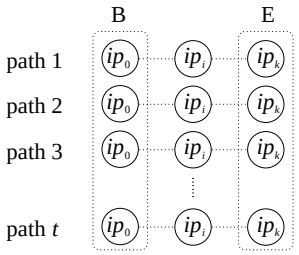


Figure 6: Proof of Lemma 5

As an example, we consider Network 2 in Figure 2, whose duplicate degree is four. The subgraph expanded by the circle of addresses $\langle 1, 2, 4, 1 \rangle$ consists of four circles. The subgraph expanded by a minimal circle of addresses $\langle 3, 4, 5 \rangle$ consists of one circle with $s = 4$: $\langle n_0, n_1, n_2, n_3, n_4, n_5, n_6, n_7, n_8, n_9, n_{10}, n_{11} \rangle$. A non-minimal circle $\langle 1, 2, 3, 4, 1 \rangle$ also expands a subgraph that has $s = 4$: $\langle n_{12}, n_{13}, n_3, n_4, n_{14}, n_{15}, n_6, n_7, n_{16}, n_{17}, n_9, n_{10}, n_{18}, n_{19}, n_0, n_1, n_{12} \rangle$.

6 Duplication Detection with Information Provided by Wireless Nature

Neighborhood views defined in last section only contain neighboring addresses and duplications can be detected in most networks except those that have special symmetrical properties. In this section, we consider definition of neighborhood views that contain more information. In Theorem 5, we observe that if neighboring addresses are consistent at all the nodes, there exists in the network a special circle that consists of patterns that have the same sequence of addresses. For example, the repeated patterns in Figure 5 are $\langle n_i^0, \ldots, n_i^k, n_{i+1}^0 \rangle$ (here we write the patterns in such a way that the first node of the next pattern is the last node of the last pattern). In order to form a circle, either the distance between two ends of each pattern is zero, which means two nodes that have address ip_0 are at the same location; or patterns do not have the same shapes and orientations, since otherwise the end of the last pattern cannot go back to the beginning of the first pattern. Since the sequence of addresses is the same for all the patterns, difference in shapes and orientations means neighbor distance information differs on nodes with the same address. Thus if neighbor distance is included in neighborhood views, inconsistency will be detected.

In practice, accurate distance information might not be available. Errors would be caused by inaccurate distance measurement or limitation in the number of bits to represent distance information. In the sequel, instead of relying on accurate distance information, our conclusion is based on inaccurate information with bounded error. We show that duplication can detected if two nodes with the same address are not too close; the meaning of "being close" is decided by the bound on the error and lengths of minimal circles in the addresses map. Note in this approach, the only modification of the original link state routing is to attach relative distances to neighbors in link state packets.

We denote the real x-coordinate (y-coordinate resp.) of node n by $x_{coor}(n)$ ($y_{coor}(n)$ resp.), and the real distance from node n to node n' in x-direction (y-direction resp.) by $dis_X(n, n')$ ($dis_Y(n, n')$ resp.). We assume each node n has distance information to each neighbor n' in x-direction and y-direction, denoted by $dis_{X\_inf}(n, n')$ and $dis_{Y\_inf}(n, n')$ respectively. Node n that has address ip generates link state packet for each of its neighbor n' that has address ip' in the form of $\langle ip, ip', d_x, d_y \rangle$, where $d_x = dis_{X\_inf}(n, n')$ and $d_y = dis_{Y\_inf}(n, n')$. Note distance information obtained by each node might differ from the real information. Let e_{rr} be the bound on distance errors defined as follows: $\forall n, \forall$neighbor n' of n, $|dis_{X\_inf}(n, n') - dis_X(n, n')| \leq e_{rr}$ and $|dis_{Y\_inf}(n, n') - dis_Y(n, n')| \leq e_{rr}$. Physical and logical neighborhood views are defined below; the term "view" is used in this section according to this definition.

Definition 5 *Given a network and a node n that has address ip, we define*

- *physical neighborhood view of $n \equiv \{\langle ip', dis_{X\_inf}(n, n'), dis_{Y\_inf}(n, n') \rangle \mid ip'$ is the address of a node n' that is physically connected n }*
- *Logical neighborhood view of $n \equiv \{\langle ip', d_x, d_y \rangle \mid \exists$ link state packet $\langle ip, ip', d_x, d_y \rangle\}$*

We now investigate the impact of distance errors on duplication detection. The following theorem shows that this approach fails only if (1) all the nodes that have the same address have the same set of neighboring addresses, **and** (2) there exists pair of nodes such that they have the same address and they are within distance $2ke_{rr}$ in both x-direction and y-direction. The intuition is that consistent views of neighboring addresses implies the existence of a circle that consists of patterns that have the same sequence of addresses, and consistent views of neighborhood implies similarity in shapes and orientations of all these patterns; thus in order to form a circle, the two ends of these patterns have to be close enough, which means two nodes with the same address are close.

Theorem 6 *Consider a network in which addresses are unique in two hops neighborhood and nodes that have the same address have the same set of neighboring addresses. At least one node has inconsistent views if any two nodes that have the same address are away at least $2k \cdot e_{rr}$ in both x-direction and y-direction, where e_{rr} is an upper bound on errors in distance information and k is the length of the special circle defined in the second part of Theorem 5.*

Proof. By Theorem 5 there is a cycle, $\langle ip_0, \ldots, ip_{k-1}, ip_0 \rangle$, in the address map such that there is a circle in the network $\langle\ n_0^0,\ n_0^1,\ \ldots,\ n_0^{k-1},\ n_1^0,\ n_1^1,\ \ldots,\ n_1^{k-1},\ \ldots,\ n_{s-1}^0,\ n_{s-1}^1,\ \ldots,\ n_{s-1}^{k-1},\ n_0^0\ \rangle$, where $s \geq 1$ and the address of n_j^i is $ip_i\ \forall j \in [0, s-1]$ (Figure 7). Assume all the nodes have consistent views. We define the following two denotations:

- $seg_{Xi} = \sum_{j=0}^{k-2} dis_X\left(n_i^j, n_i^{j+1}\right) + dis_X\left(n_i^{k-1}, n_{(i+1)\%s}^0\right)$

  This is the real distance in x-direction from n_i^0 to $n_{(i+1)\%s}^0$.

- $seg_{X\_inf} = \sum_{j=0}^{k-2} dis_{X\_inf}\left(n_i^j, n_i^{j+1}\right) + dis_{X\_inf}\left(n_i^{k-1}, n_{(i+1)\%s}^0\right)$

  Note the value of seg_{X_inf} does not depend on i, because for all i, $dis_{X\_inf}(n_i^j, n_i^{j+1})$ has the same value since $n_i^j = ip_j$ and $n_i^{j+1} = ip_{j+1}$ and views are consistent on all the nodes.

By the definition of e_{rr}, we have $|seg_{Xi} - seg_{X\_inf}| \leq ke_{rr}$. Since $\sum_{i=0}^{s-1} seg_{Xi} = 0$, we have $seg_{X\_inf} \in [-ke_d, ke_d]$, that is, $seq_{Xi} \in [-2ke_d, 2ke_d]$. So we prove nodes n_i^0 and $n_{(i+1)\%s}^0$ are within x-distance $2ke_{rr}$. Similarly, we can prove nodes n_i^0 and $n_{(i+1)\%s}^0$ are within y-distance $2ke_{rr}$. Thus there exists two nodes with the same address that are within distance $2ke_{rr}$ in both x-direction and y-direction. Contradiction!

■

Now we discuss how nodes decide the number of bits to represent the number of distance information. We consider a network in which transmission range of nodes is R. Letting d_x (d_y resp.) be the distance within any two nodes that have the same address in x-direction (y-direction resp.), we have $max\{|d_x|, |d_y|\} \geq \frac{R}{\sqrt{2}}$ by the assumption that addresses are unique within in two hops neighborhood. In order to detect duplication, we require $2ke_{rr} \leq \frac{R}{\sqrt{2}}$,

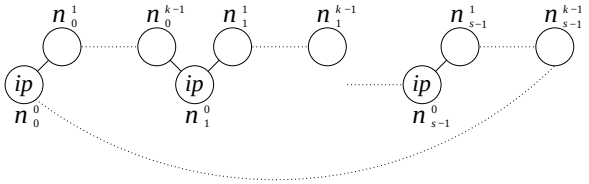


Figure 7: Proof of Lemma 6

that is, $e_{rr} \leq \frac{R}{2\sqrt{2}k}$. If b bits are used to represent accurate distance in link state packets, we have $e_{rr} \leq \frac{R}{2^b}$. So all duplications can be detected if $\frac{R}{2^b} \leq \frac{R}{2\sqrt{2}k}$, that is, $b \geq 1.5 \log k$. Nodes can get an upper bound on k as the maximum length of minimal circles in the addresses map; the the "minimal" property of such a circle shown in Theorem 5 implies high possibility of a small k. Note a trivial upper bound on k is the number of *addresses*, which is smaller than a usual bound that depends on the number of nodes or the length of some assumed global unique keys.

7 Conclusion

We investigated duplicate address detection under the assumption that addresses are unique within two hops neighborhood. We propose two definitions of neighborhood views and duplication detection is done by comparing the physical and logical neighborhood views of each node. We show traditional neighborhood information, neighboring addresses, is not sufficient to detect duplicate address while duplication can be detected by using neighbor distance information, which is available in ad hoc networks due to its wireless nature.

References

[1] A.Benlarbi, J. Cousin, R. Ringot, A. Mamouni, and Y. Leroy. Interferometric positioning systems by microwaves. In *Proc. Microwaves Symp.*, Tetuan, Morocco, 2000.

[2] P. Boldi and S. Vigna. Universal dynamic synchronous self-stabilization. *Distributed Computing*, 15-3:137–153, 2002.

[3] S. Boudjit, A. Laouiti, P. Muhlethaler, and C. Adjih. Duplicate address detection and autoconfiguration in olsr. In *SNPD-SAWN '05*, pages 403–410, Washington, DC, USA, 2005. IEEE Computer Society.

[4] J. Cartigny, D. Simplot, and I. Stojmenovic. Localized minimum-energy broadcasting in ad-hoc networks. In *INFOCOM 2003*, 2003.

[5] B. S. Chlebus, L. Gasieniec, A. Ostlin, and J. M. Robson. Deterministic radio broadcasting. In *Automata, Languages and Programming*, pages 717–728, 2000.

[6] A. E. F. Clementi, A. Monti, and R. Silvestri. Selective families, superimposed codes, and broadcasting on unknown radio networks. In *Proc. 12th Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 709–718, Washington, DC, 2001.

[7] R. Droms. Dynamic host configuration protocol, 1997.

[8] T. K. M. Yamashita. Computing on anonymous networks: Part i — characterizing the solvable cases. *IEEE Transactions on Parallel and Distributed Systems*, 7(1):69–89, 1998.

[9] T. K. M. Yamashita. Computing on anonymous networks: Part i — decision and membership problems. *IEEE Transactions on Parallel and Distributed Systems*, 7(1):90–96, 1998.

[10] N. Mitton, E. Fleury, I. GuÃ©rin-Lassous, B. SÃ©ricola, and S. Tixeuil. On fast randomized colorings in sensor networks. technical report LRI-1416, INRIA, Jun. 2005.

[11] M. Mohsin and R. Prakash. An ip address configuration algorithm for zeroconf mobile multihop ad hoc networks. In *Int'l. Wksp. Broadband Wireless Ad Hoc Networs and Services*, Sept 2002.

[12] S. Nesargi and R.Prakash. Manetconf: Configuration of hosts in a mobile ad hoc network. In *INFOCOM 2002*, June 2002.

[13] S. Ni, Y. Tseng, Y. Chen, and J. Sheu. The broadcast storm problem in a mobile ad hoc network. In *Proc. 5th Annual ACM/IEEE International Conference on Mobile Computing and Networking*, pages 151–162, Seattle, WA, 1999.

[14] R. C. Shah and J. M. Rabaey. Energy aware routing for low energy ad hoc sensor networks. In *IEEE WCNC*, 2002.

[15] S. Thomoson and T. Narten. Ipv6 stateless address autoconfiguration. *RFC 2462*, Dec 1998.

[16] N. H. Vaidya. Weak duplicate address detection in mobile ad hoc networks. In *Proc. 3rd ACM International Symposium on Mobile Ad Hoc Networking and Computing*, pages 206–216. ACM Press, 2002.

[17] K. Weniger. Passive duplicate address detection in mobile ad hoc networks. In *IEEE WCNC*, New Orleans, LA, Mar 2003.

[18] K. Weniger. Pacman: Passive autoconfiguration for mobile ad hoc networks. *IEEE Journal on Selected Areas in Communications*, 23(3):507–519, 2005.

[19] E. Wesel. *Wireless Multimedia Communications: Networking Video, Voice, and Data*. Addison-Wesley, Reading, MA, 1998.

Unité de recherche INRIA Sophia Antipolis
2004, route des Lucioles - BP 93 - 06902 Sophia Antipolis Cedex (France)

Unité de recherche INRIA Futurs : Parc Club Orsay Université - ZAC des Vignes
4, rue Jacques Monod - 91893 ORSAY Cedex (France)
Unité de recherche INRIA Lorraine : LORIA, Technopôle de Nancy-Brabois - Campus scientifique
615, rue du Jardin Botanique - BP 101 - 54602 Villers-lès-Nancy Cedex (France)
Unité de recherche INRIA Rennes : IRISA, Campus universitaire de Beaulieu - 35042 Rennes Cedex (France)
Unité de recherche INRIA Rhône-Alpes : 655, avenue de l'Europe - 38334 Montbonnot Saint-Ismier (France)
Unité de recherche INRIA Rocquencourt : Domaine de Voluceau - Rocquencourt - BP 105 - 78153 Le Chesnay Cedex (France)

Éditeur
INRIA - Domaine de Voluceau - Rocquencourt, BP 105 - 78153 Le Chesnay Cedex (France)
http://www.inria.fr
ISSN 0249-6399