Formal Verification of a Multicast Protocol in Mobile Networks

M. Matash Borujerdi, S.M. Mirzababaei

Abstract—As computer network technology becomes increasingly complex, it becomes necessary to place greater requirements on the validity of developing standards and the resulting technology. Communication networks are based on large amounts of protocols. The validity of these protocols have to be proved either individually or in an integral fashion. One strategy for achieving this is to apply the growing field of formal methods. Formal methods research defines systems in high order logic so that automated reasoning can be applied for verification. In this research we represent and implement a formerly announced multicast protocol in Prolog language so that certain properties of the protocol can be verified. It is shown that by using this approach some minor faults in the protocol were found and repaired. Describing the protocol as well-formed formulas. J. Holzman [1] categorized the methods in 94, 95 and divided the manual methods to static methods and formal methods. M. Goda [6,7] categorized these methods are used to validate network protocols, manual methods and formal methods. M. Goda [6,7] categorized these methods in 94, 95 and divided the manual methods to static and well-formed formulas. J. Holzman [1] categorized the automatic methods into four groups: full search, controlled partial search, random simulation, and formal verification. There are other similar works in formal verification of IPv6 partial search, random simulation, and formal verification.

Keywords— Formal methods, MobiCast, Mobile Network, Multicast.

I. INTRODUCTION

NOWADAYS as the computer systems become increasingly complex, the difficulty of developing highly reliable technology increases as well. Numerous design flaws have been chronicled such as the Intel Pentium processor floating-point bug and the Denver Airport baggage handling system problem. System developers have to test and verify the products before releasing them to the market to avoid later problems. One of the most complex system categories that may have flaws is the computer network. Two major types of methods are used to validate network protocols, manual methods and formal methods. M. Goda [6,7] categorized these methods in 94, 95 and divided the manual methods to static and well-formed formulas. J. Holzman [1] categorized the automatic methods into four groups: full search, controlled partial search, random simulation, and formal verification. There are other similar works in formal verification of IPv6 [2] and also the session layer [5]. The potentials of formal methods to provide solutions in these areas are also found in [8, 9, and 10].

This paper demonstrates how a formerly announced network protocol such as MobiCast can be specified for theorem proving in Prolog language. Then semi-automated reasoning can be applied to the specification to verify certain properties. In the following sections we will first discuss the MobiCast protocol. Then we will show the relationships between definitions of graphs and communications and the implementation of the MobiCast. After that we will discuss our verification method before our final conclusions.

II. MobiCAST

MobiCast [11] introduced by Lin Tan is an appropriate protocol for a network with micro-cells whose base-stations stand on a high speed wired network capable of transmission of multicasting [3] packets to several recipients [12, 13]. This method isolates the movement of mobile hosts from multicast tree to minimize disconnections during sessions. This mechanism uses the hierarchical mobility management to isolate the movement of MHS(Mobile Host) from the main multicast transfer tree. Every external domain must use an agent.

To send a multicast packet the MH has to encapsulate it in a regular packet and send it to domain agent. The agent opens the packet and sends it to receivers, like a multicast packet. Whenever a station wants to subscribe in a multicast group it must send its subscription request to DA (Domain Agent) via its BS (Base Station). The DA subscribes to that multicast group instead of the MH and receives the requested packets and then sends them to the MH. There is another multicast address that denoted as translated multicast address. This translated multicast address has to be unique in the domain and must be related to main multicast group. The BS subscribes for reception of these packets in the translated multicast group (whose address is the translated one), and sends the received packets to the MH. The roaming of the mobile hosts is transparent from other group members while it exists in the domain of the mentioned DFA (Domain Foreign Agent) and there is no need for new calculation of the multicast main tree resulting in minimizing the breaks in multicast sessions.

The Lin Tan's protocol organizes and builds bigger cells with composition of micro-cells and names it DVM (Dynamic Macro-Cell). When an MH subscribes to a multicast group via DFA the current base station will inform the other base

Manuscript received January 9, 2004.

M. Matash Borujerdi, is with the Department of IT and Computer Engineering, Amirkabir University of Technology, Tehran, Iran, (e-mail: borujerdii@aut.ac.ir).

S.M. Mirzababaei, also is with the Department of IT and Computer Engineering, Amirkabir University of Technology, Tehran, Iran, (e-mail: mirzababaei@morva.net).
stations that are in its DVM to subscribe in that translated multicast group. While only the serving base station actively forwards multicast data packets to the mobile host, the other base stations in the same DVM buffer recent packets and quickly forwards those to the mobile host whenever a handoff occurs. This provides short handoff latency, and the use of buffers at the base stations reduces packet loss due to handoff. It also eliminates multicast group join and graft latencies since the new base station has already subscribed to the multicast group prior to the handoff. Hence, the disruptions to the multicast session due to handoffs of mobile host group members are minimized.

Each multicast group is associated with a translated multicast group address, and serving BSs of interested MHs only need to subscribe to this translated multicast group to receive the desired multicast data. Besides delivering multicast data in an efficient manner (as compared to multiple unicasts to interested MHs within its domain), the use of multicast as the forwarding mechanism from the DFA to interested MHs in its domain also alleviates the DFA from the task of keeping track of the exact location of the MH to ensure correct multicast data delivery. Since physically adjacent cells are most likely to reside on the same network segment, the extra network load generated due to the other member BSs in the same DVM subscribing to the same multicast group is negligible. This is especially so for the case of shared medium networks such as Ethernet.

MobiCast is developed to work with IP and is compatible with existing multicast routing algorithms such as DVMRP, CBT, MOSPF, PIM-D and PIM-SM. The base stations are network-layer routers with buffers, and are capable of subscribing to multicast groups. Compared to a link-layer solution adopted by most wireless LAN, a network-layer base station is capable of forwarding only those multicast packets with interested mobile receivers in its cell, thus achieving efficient utilization of wireless bandwidth. Furthermore, a network-layer base station is able to differentiate packets with different service types for IPv6 so as to support QoS (Quality of Service) for mobile hosts in its cell. This scheme aims to support best effort IP multicast efficiently for mobile hosts in an environment with Micro-cells, while maintaining the quality of the multicast session during handoffs.

As appears in figure 1 when each MH moves, it leaves the wireless coverage of one cell and enters into another, resulting in a handoff in the base stations. In small wireless cells at the fringes of the Internet, such handoffs during a multicast session will be frequent as wireless cells may be of the size of a few meters. When an MH arrives at a foreign network and obtains an in-care-of address, the MH sends a location update message to inform it’s HA (Home Agent) of its care-of address. The care-of address identifies an FA (Foreign Agent) in the foreign subnet where the MH is. The FA can be a separate node or reside in the MH. The MH has to subscribe to the desired multicast group via FA1 when it is at subnet A and via FA2 when it moves to subnet B.

MobiCast has the following advantages. First, the use of hierarchical mobility management architecture separates the mobility of the MH from the main multicast delivery tree. As long as the MH remains within the domain of the DFA, the mobility of the MH is shielded from the rest of the multicast group. No re-computation of the main multicast delivery tree is needed. Second, this scheme requires MHs, which are interested in receiving multicast data to re-subscribe again via the DFA when they are in a foreign domain. This approach is somewhat similar to the remote subscription method proposed by Mobile IP, and network routes taken by the multicast packets to the MHs are efficient and the inefficiencies and scalability problems associated with the Mobile IP are totally avoided. Third, this scheme uses multicast to forward the multicast packets from the DFA to the interested MHs within its domain.
Another method is to represent the whole graph as one data object. A graph can be thus represented as a pair of two sets: nodes and edges. Each set can be represented as a list; each edge is a pair of nodes. Let us choose the functor graph to combine both sets into a pair, and the functor e for edges. Then one way to represent the (undirected) graph in the following code:

\[
G1 = \text{graph}([a,b,c,d],[e(a,b),e(b,d),e(b,c),e(c,d)]).
\]

To represent a directed graph we can choose the functor digraph and a (for arcs). The directed graph of the following code is then:

\[
G2 = \text{digraph}([s,t,u,v],[a(s,t,3),a(t,v,1),a(t,u,5),a(u,t,2),a(v,u,2)]).
\]

If each node is connected to at least one other node then we can omit the list of nodes from the representation as the set of nodes is then implicitly specified by the list of edges. Yet another method is to associate with each node a list of nodes that are adjacent to that node. Then a graph is a list of pairs consisting of a node plus its adjacency list. Our example graphs can then, for example, be represented by:

\[
G3 = [p(a,[arcto(b,3),arcto(c,1)]),p(b,[arcto(e,2),arcto(f,1)])].
\]

What will be the most suitable representation will depend on the task to be performed on graphs. Two typical operations are finding a path between two given nodes and finding a subgraph with some specified properties. Finding a spanning tree of a graph is an example of the latter operation. To find a Steiner tree we must define these rules: The spanning tree is a connected sub-graph of the main graph that has no cycles. To find the minimum spanning tree we start from an empty set of edges, and add more edges in a way that no cycles be generated and continue to add edges until no more edges can be added. This set will be the spanning tree. And the following is also a useful rule in accomplishing the task: We can add those edges that have only one terminal in the set. Following prolog rules is the implementation of the above rules:

\[
\text{stree}(G,T) :- \text{member}(E,G), \text{spread}([E],T,G).
\]

\[
\text{spread}(T1,T,G) :- \text{addedge}(T1,T2,G), \text{spread}(T2,T,G).
\]

\[
\text{addedge}(T,[e(A,B)|T]) :- \text{adjacent}(A,B,G), \text{node}(A,T), \text{not}(\text{node}(B,T)).
\]

\[
\text{adjacent}(N1,N2,G) :- \text{member}(e(N1,N2),G); \text{member}(e(N2,N1),G).
\]

\[
\text{node}(N,G) :- \text{adjacent}(N,_,G).
\]

All three arguments of the spread (T1,T,G) are sets of the edges. G is fully connected and T1 and T are sub-sets from G that represents the tree. T is a spanning tree that built form adding zero or one edge from G to T1 in every step. We define the above rules in declarative manner as follows.

\[
\text{stree}(G,T) :- \text{subset}(G,T), \text{tree}(T), \text{covers}(T,G).
\]

\[
\text{tree}(T) :- \text{connected}(T), \text{not}(\text{hasacycle}(T)).
\]

\[
\text{connected}(G) :- \text{node}(A,G), \text{node}(B,G), \text{not}(\text{path}(A,B,G,\_)).
\]

\[
\text{hasacycle}(G) :- \text{adjacent}(N1,N2,G), \text{path}(N1,N2,G,[N1,X,Y]_{\_}).
\]

\[
\text{covers}(T,G) :- \text{node}(N,G), \text{not}(\text{node}(N,T)).
\]

\[
\text{subset}([],\_). \text{subset}([X|S],S) :- \text{subset}(S,Ss). \text{subset}([X|S],[X]|Ss) :- \text{subset}(S,Ss).
\]

Now we must choose some special nodes and use them to find a Steiner tree, so the covers rule has to use from G1 instead of G. To define the MobiCast protocol we must also define some high level definitions and represent them in prolog.
1- Building a DVM: we must build a DVM with the cell and its adjacent cells, when a mobile host enters a cell then destroy the DVM when the mobile host leaves the cell. This work will be done via subscribing to a new multicasting group.

2- Subscribing to the multicasting group: the mobile host instructs the nearest DFA to subscribe to the interested multicast group.

3- Unsubscribing from the multicasting group: the mobile host instructs the nearest DFA that this node is not interested to be subscribed to the multicasting group any more and the DFA can leave the multicasting group according to its policies.

4- Subscription of the DFA in the multicast group: DFA subscribes to the multicast group, so the tree has to be calculated again.

5- DFA Unsubscribing from multicast group: DFA leaves the multicast group; this event can lead to a new tree calculation.

6- Sending a packet to the multicast group: The mobile host sends a packet to the DFA so it can send the packet to the multicast group instead of the sender.

7- Receiving a packet by DFA: When the DFA receives a packet from a translated multicast address; it converts the translated multicast address to the main multicast address and resends it. Otherwise if the received packet has the main multicast address, the translated address has to be found to forward the packet to the DVM. These steps have been show in the figure 2.

As we only need to prove the correctness of the protocol and do not need to give good performance, we can simplify the search for the multicast addresses or translated multicast addresses by defining the multicast addresses as those above 1024 and the translated multicast addresses as those that are below 1024. This way the MobiCast is represented and implemented in Prolog as illustrated below.

```
database
single receivedpacket(node, integer)
groupNo(graph, integer)
mhDFA(mhid, node)
mhGroup(mhid, integer)
cell(mhid, node)
```

```
classes
carrierNow(CurNode, [], X):- !.
carrierNow(CurNode, [e(CurNode, Node)|_], Node).
carrierNow(CurNode, [e(Node, CurNode)|_], Node).
carrierNow(CurNode, [_|T], Node):- carrierNow(CurNode, T, Node).
carrierInDFA(CurNode, X, I):- I>1024, II=I-1024, groupNo(T, II), carrierNow(CurNode, T, X).
carrierInDFA(CurNode, X, I):- I<=1024, II=I+1024, groupNo(T, II), carrierNow(CurNode, T, X).
sendPacket(MHId, X):- mhDFA(MHId, DFA), mhGroup(MHId, I), carrierInDFA(DFA, X, I).
enMemberVAddr(MHId, I):- mhDFA(MHId, DFA), cell(MH, BTS), constructDVM(BTS, I), enMemberAddr(DFA, I), II=I-1024, assert(groupNo(MHId, II)).
enMemberAddr(DFA, I):- groupNo(T, I), groupNo(G, I), addToSubGraph(DFA, G, G1), stree(G1, T1), retract(groupNo(T, I)), retract(groupNo(G, I)), assert(groupNo(T1, I)), assert(groupNo(G1, I)).
deMemberVAddr(MHId, I):- mhDFA(MHId, DFA), deMemberAddr(DFA, I), cell(MH, BTS), destructDVM(BTS, I), retract(mhGroup(MHId, I)), retract(groupNo(T, I)), groupNo(G, I), OmitFromGraph(DFA, G, G1), stree(G1, T1), retract(groupNo(T, I)), retract(groupNo(G, I)), assert(groupNo(T1, I)), assert(groupNo(G1, I)).
makeDVM(NBts, OBts, I):- destructDVM(OBts, I), constructDVM(NBts, I).
```

Figure 2: The protocol stages dependencies
The program analysis studies the relation between the input and the output of a program. Formally, a program consists of an input (variable) vector \( \mathbf{x} = (x_1, \ldots, x_m) \), a program (variable) vector \( \mathbf{y} = (y_1, \ldots, y_n) \), and an output (variable) vector \( \mathbf{z} = (z_1, \ldots, z_k) \), and a finite direct graph \((V,A)\) such that the following conditions are satisfied.

1- In the graph \((V,A)\), there is exactly one vertex, called the start vertex \( S \in V \). That is not a terminal vertex of any arc; there is exactly one vertex called the halt vertex \( H \in V \), that is not an initial vertex of any arc; and every vertex \( v \) is on some path from \( S \) to \( H \).

2- In \((V,A)\), each arc \( a \) not entering \( H \) is associated with a qualifier-free formula \( P_a(\mathbf{x},\mathbf{y}) \) and an assignment \( \gamma \leftarrow f_a(\mathbf{x},\mathbf{y}) \); each arc entering the halt vertex \( H \) is associated with a qualifier-free formula \( P_a(\mathbf{x},\mathbf{y}) \) and an assignment \( \gamma \leftarrow f_a(\mathbf{x},\mathbf{y}) \). (The \( P_a \) is called the testing predicate associated with arc \( a \) and \( P_a(\mathbf{x},\mathbf{y}) \) is called the testing formula associated with arc \( a \))

3- For each vertex \( (v \neq H) \), let \( a_1, a_2, \ldots, a_n \) be all the arcs leaving \( v \) and let \( P_{a_1}, P_{a_2}, \ldots, P_{a_n} \) be the testing predicates associated with arcs \( a_1, a_2, \ldots, a_n \), respectively. Then for all \( \mathbf{x}, \mathbf{y}, \mathbf{z} \), one and only one of \( P_{a_1}(\mathbf{x},\mathbf{y}), P_{a_2}(\mathbf{x},\mathbf{y}), \ldots, P_{a_n}(\mathbf{x},\mathbf{y}) \) is True.

The formal representation is used to provide answers to following problems:

1- Termination problem: Given a certain input, will this program terminate?

2- Response problem: Given a certain input, if the program terminates, what is the output of the program?

3- Correctness problem: Given a certain input, will the output of this program satisfy the specification (input-output relationship) of the program?

4- Equivalence problem: Given two programs, will the programs yield the same results if the inputs are the same?

5- Specialization problem: Given a program \( P \) that is written to accept a set \( S \) of inputs, if we are only interested in a nonempty subset \( S^* \) of \( S \), how can we simplify \( P \) to another \( P^* \) such that \( P^* \) runs faster on \( S^* \) than \( P \) does?

Evidently, we need some information to answer any of the above questions. In general, we need the following information:

1- Axioms describing the execution of the program. We describe this by \( A_1 \).

2- Axioms concerning testing predicates and assignment functions. For example, we might need the axioms concerning equality or some appropriate induction schema. We describe this by \( A_2 \).

3- Axioms concerning the input. For instance, we might require the input to be positive integers. We describe this by \( A_3 \).

We assume that \( A_1 \land A_2 \land A_3 \) is consistent. But, we have to prove that \( A_1 \land A_2 \land A_3 \) is consistent with the description of the output of the program. There is a theorem that says: Given a program \( P \), let \( S \) denote the set of clauses representing \( A_1 \land A_2 \land A_3 \), Then \( S \) is satisfiable[14].

So the declarative information of the MobiCast protocol prepared using Prolog language. Of course the non-declarative features in the Prolog language such as the depth-first search rule, a non-logical negation (by failure), and a number of non-logical operations such as the test predicate (e.g., var) and the cut are nevertheless necessary to make Prolog reasonably efficient. Prolog language has some first order logic predicates that formed in clause forms (in CNF composition and without quantifiers in form of prenex like). We used the SLD resolution for automatic theorem proving that showed below.

**domains**

expr = symbol; if (expr, expr);
and (expr, expr); or (expr, expr); not (expr); true

database

nondeterm clause (expr)
done (expr, expr, expr)
modified

predicates

nondeterm contradiction
nondeterm remove_true_clause
nondeterm simplify_clause
nondeterm resolution_step
nondeterm delete (expr, expr, expr)
nondeterm in(expr, expr)
translate(expr)
translate_neg(expr)
nondeterm transform(expr, expr)
nondeterm run

clauses

clause(true).
contradiction :- contradiction(X), contradiction(not(X)), write("contradiction found(Formula is true). ").
remove_true_clause :- clause(C), in(P, C), in(not_(P), C), retract(clause(C)).
simplify_clause :- clause(C), delete (P, C, C1), in(P, C1), retract(clause(C)), assert(clause(C1)).
resolution_step :- clause(P), clause(C), delete (not_(P), C), not(done(P, C, P)), assert(clause(C1)), assert(done(P, C, P)).
resolution_step :-

clause(not_(P)), clause(C), delete_(P,C,C1), not(done(not_(P), C,P)), assert(clause(C)), assert(done(not_(P), C,P)).

resolution_step:-
clause(C), delete_(P,C1,CA), clause(C2), delete_(not_(P), C2, CB), not(done(C1,C2,P)), assert(clause(or_(CA,CB))), assert(done(C1,C2,P)).

% delete(P,E,E1) if deleting a disjunctive subexpression P from E gives E1
delete_(X,or_(X,Y),Y).
delete_(X,or_(Y,X),Y).
delete_(X,or_(Y,Z),or_(Y,Z1)):-delete_(X,Z,Z1).
delete_(X,or_(Y,Z),or_(Y1,Z)):-delete_(X,Y,Y1).

% in(P,E) if P is a disjunctive subexpression in int(X,Y).
in(X,Y):-delete_(X,Y,). % double negation

% transform(Formula) : translate propositional Formula into clauses and assert each resulting clause C as clause(C)
translate(and_(F,G)):-!,translate(F),translate(G).
translate(F):-transform(F,NewF),!,translate(NewF).
translate(F):-assert(clause(F)).

% transform subexpression
transform(or_(X,Y),or_(X1,Y)):-transform(Y,Y1).
transform(not_(X),not_(X1)):-transform(X,X1).

% Test Stub
run:-contradiction.
run:-remove_true_clause, assert(modified), fail.
run:-simplify_clause, assert(modified), fail.
run:-resolution_step, assert(modified), fail.
run:-modified, retract(modified), run.
translate(not_(X)): translate(not_(X)).

For example, we now show the abstract level verification of two parts of the program. First we verify the tree building in the program and then we will verify packet delivery mechanism. At the first step we must have a description of the program execution, statements, inputs and outputs of tree building. There is no need to isolate these parts from each other so we extracted all of these descriptions manually from the prolog code. We need an automatic converter if we want a full automatic verifier

The database

SrcNodes(node)
DestNodes(node)
Arc(node,node)
TreeArcs(node,node)

The clauses
Tree→SrcNodes(S) ∧ DestNodes(D) ∧ Arc(S,D) ∧ TreeArcs(S,D) ∧ SrcNodes(D).
TreeArcsDesc(N1,N2) → ~Path(N2,N1).
Path(N2,N1) → TreeArcs(N2,N3) ∧ Path(N3,N1).

Now we feed the resolution program with these facts and the negation of the result. These clauses will be found in memory in case of program trace.

Path(b,a)  ~Path(b,a)  Arc(a,b)
TreeArcs(a,b)  ~TreeArcs(N,b)

And then the program concludes the contradiction that means the consistency of the program. Again when we want to verify the packet delivery mechanism, we have to do as the above, so we prepare the description of the program execution, statements, inputs and outputs.

The database
MhDFA(mhid,node)
MhGroup(mhid,grpid)
GroupNo(tree,grpid)
ReceivedPacket(node,mhid)
NextNode(node,tree,node)
TransG(grpid,grpid)

The clauses
SendPacket(S,M)→MhDFA(S,D) ∧ MhGroup(S,I) ∧ CarryInDFA(D,M,I).
CarryInDFA(D,M,I)→TransG(I,II) ∧ GroupNo(T,II) ∧ CarryInTree(T,D,M).
CarryInTree(D,T,M)→ReceivedPacket(D,M) ∧ NextNode(D,T,Dx) ∧ CarryInTree(Dx,T,M).

Then we present these sentences to the SLD resolution program, with the negation of the result "not(ReceivedPacket(N,M))". These clauses will be found in memory in case of program trace.

SendPacket(a,b)  ~ReceivedPacket(c,b)
MhDFA(a,c)  MhGroup(a,d)
CarryInDFA(c,b,d)  TransG(d,e)
CarryInTree(c,f,b)  ReceivedPacket(c,b)
NextNode(c,f,b)

So then the program concludes the contradiction between the second (negation of the result) and 8th fact that means the consistency of the program.

V. CONCLUSIONS

The specification of the MobiCast protocol partially presented here demonstrates how the designers of a protocol
can work out inconsistencies prior to release. Several inconsistencies were found in the protocol and resolved with minor variation, though mostly with contradictions between different sections in the protocol. The major benefit of this work is that the result is a well-defined set of criteria that implementers must meet. As such, it is advocated that formal methods should become a critical component of the standardization process. In fact, if an implementation is first defined in formal languages and then in the final language, the implementation can be verified against the specification, increasing the probability that different implementations would have no inconsistencies. Future work will also include verifying that two different protocols can co-exist together such as TCP-IP [4].

REFERENCES