Privacy in the Pervasive Era: A Distributed Firewall Approach

Leonardo Maccari
DISI – University of Trento
Trento, Italy
Email: leonardo.maccari@disi.unitn.it

Renato Lo Cigno
DISI – University of Trento
Trento, Italy
Email: locigno@disi.unitn.it

***Abstract*—Pervasive computing and communications are (slowly) enabling local ad-hoc services. Preserving privacy in a pervasive environment is one of the key challenges ahead: How can users define their "communication boundaries"? how can the network avoid wasting resources and eventually collapse under the burden of undesired traffic that will be discarded at the receiver machine? In this paper we propose the adoption of distributed filtering techniques implementing a network-wide firewall whose goal is defining precisely, and under the user control, the boundaries in space, time, information content, and logical addressing of a user communication scope. Initial results based on an implementation integrated with OLSR are presented.**

***Index Terms*—Privacy; Distributed Filtering; Firewalling; Ad Hoc Networks; Mesh Networks; Local Social Networks.**

I. Introduction

With the success of the Internet and its convergence with mobile communications, people lives take place (also) in the cyber-$\alpha\gamma o\rho\acute{\alpha}$[1]. Life in public spaces is governed by behavioral rules, respect of other people space and privacy, and it is traditionally limited in time and space: retreating home or changing place changes the situation and the rules. The cyber-$\alpha\gamma o\rho\acute{\alpha}$ instead is anywhere at anytime, collapsing a multiplicity of situations and scenarios into a single social space that imposes new rules of behavior: rules that are often difficult to understand, more difficult to describe, and even more difficult to enforce.

Fig. 1 depicts a high level view of the cyber-$\alpha\gamma o\rho\acute{\alpha}$: Users have direct connections within a mesh network, interact one another and with ubiquitous sensors and other machines (including cars) with communication capabilities. Mesh nodes may act as gateways to the global Internet, user nodes can be connected across the mesh and, at the same time to the global mobile communication system. What messages should a user receive? who and what are the legitimate destinations of, e.g., a user location? what is the number of hops in the ad-hoc network that defines the user *circle* (or *aspect*, using a social-network derived terminology to define a limited set of contacts)?. These are all legitimate questions which require the definition of filtering rules, their diffusion in the ad-hoc network and their enforcement in some (all?) nodes.

[1] $\alpha\gamma o\rho\acute{\alpha}$ is the original Greek spelling of 'agora', the public open space of political, social, and commercial life in ancient Greek city-states.

Along with user-centric policy definitions, other factors opening more questions are performance and security. Where is the best place to enforce a rule? Is there a possibility to enforce a network-wide policy in a cooperative way? is it possible to reduce the impact of unwanted/malicious traffic?

To the best of our knowledge in a pervasive environment this is a novel problem, which has so far never been tackled in its whole complexity. In this paper we describe a scenario suitable to analyze this problem and introduce our initial efforts to manage one of its sub-problems[2].

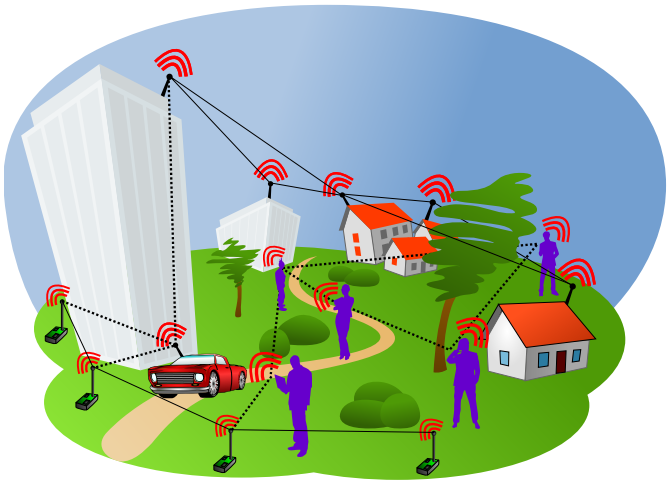

Figure 1. Nomadic users in a general mesh network scenario

A. Considered Scenario

Consider again the scenario defined in Fig. 1. A mesh network owned by a manager (public or private) guarantees connectivity on a wide area. Users with their terminals extend the mesh with ad-hoc capabilities, they may communicate directly if they are in range, they can relay traffic for their neighbors creating multi-hop paths and they can use the mesh network to extend their reachability over a larger area. As a possible traffic source we choose a very popular application, social networking, that can generate various types of traffic. Imagine then, that the network extends in an ad-hoc fashion

[2] The interested reader can find an extended version of this paper including more analysis and details at the url: http://disi.unitn.it/locigno/preprints/TR-DISI-11-481.pdf

a web-based social network, so nodes may share some application parameters but that most of them do not have Internet connectivity. Some of the issues of interest are:

1) User A wants to limit his visibility to a specific *circle*, i.e., a certain range of hops around him, or, if nodes are geolocalized, a limited area where he resides;
2) User B wants to be globally visible, but wants some of his traffic not to travel outside a specific circle;
3) The network manager wants to forbid a certain traffic type;
4) Node C wants to query node E for a specific service (e.g. friendship request), but node E is not available for that service. It is the interest of the network that the request of node C is discarded, or stopped before it wastes network resources.

While the first two issues deal mostly with users' privacy, the others deal with security and robustness of the network. Note that the presence of a managed mesh network is optional, in cases where the network is purely ad-hoc the third issue does not apply.

Three problems must be resolved: *i)* each node should define a set of rules $\mathcal{S}_i$ for traffic involving himself; *ii)* an efficient way of distributing $\mathcal{S}_i$ must be found; and *iii)* a proper strategy for the implementation of rule-sets must be devised, which achieves some given goals for the benefit of both the network and the users.

This paper focus on the third issue: in a social network with $\mathcal{N}$ nodes we assume that the software application contains various pre-defined rule-sets $\mathcal{S}_i$ identified by some IDs so that nodes need to share only the selected IDs using the ad-hoc network, and not the whole $\mathcal{S}_i$. The union of the single rule-sets define the global rule-set of the network: $\mathcal{S} = \bigcup_{n_i \in \mathcal{N}} \mathcal{S}i$

The problem of rule-set implementation can be mapped on a problem of *distributed firewall*, i.e., a system where the rule-sets $\mathcal{S}_i$ are distributed among the nodes of the network, which in turn cooperate by enforcing them to the best of their capabilities.

In the general case the cost of analyzing the rule-set is linear in the set dimension and meaningful scenarios include hundreds of nodes each one with tens to hundreds of desired rules, so that there are thousands of rules to check for each packet at each hop. This makes the trivial solution of enforcing the entire set $\mathcal{S}$ in any node normally unfeasible. The problem we tackle is what subset of $\mathcal{S}$ each node will enforce, and how to evaluate the chosen approach.

By contrast more performing filtering algorithms as the ones described in [1] may require dedicated hardware and are generally suitable for quasi-static rule-sets, since they keep complex data structures. For this reason their application to our scenario would be just as challenging as applying standard algorithms.

II. Firewalling in ad-hoc networks

A firewall is a network host or router that implements certain actions on packets described by a set of rules. A rule is defined as a pattern matching part and the corresponding action. Firewalls can be stateless (the decision on each packet is independent) or stateful, in this paper we concentrate on stateless firewalling. We will use Netfilter/IPtables terminology in the rest of the paper to describe firewalls. Netfilter/IPtables is the software that implements the firewall in kernel and user space of Linux, which is the base of the majority of commercial firewalls and of many mobile applications. It implements a linear list of rules matching fields of a packet and a target that can be an action such as drop, accept, mangle. Functionally, it divides the point where the filter is applied in *chains*, the input chain is used for packets that are destined to the firewall, the forward chain is applied to packets that are going to be routed through the firewall and the output chain is for packets generated from the firewall itself (see [2] for details). In our simulator the output and forward chains have been merged for simplicity (they behave like the postrouting chain in Linux).

The rules may match a UDP/TCP port or an application level packet content, for simplicity we consider different services as bound to a specific port (chat, presence, location queries, ...) but this approach can be moved to upper layers in the stack to verify other properties of packets, for instance, identify the references to a certain IP address into routing messages.

An initial model of distributed firewall has been proposed by Bellovin et al. in [3] where the firewall is delocalized from a bastion host to the endpoints of a still traditional centralized network. Recently, the subject has received more attention from some authors that approached the possibility of performing a true distributed firewalling. Works like [4]–[6] introduce the concepts of distributed filtering and study the performance and initial integration with routing strategies. Other authors, see [7], [8], reverse the problem and propose networks in which communication is possible only if a previous security handshake has been performed end-to-end (deny-by-default networks). In this paper we follow the first approach, concentrating on the enforcement of the rules and the integration with proactive routing.

III. Initial design and results

In this section we report some initial results on the implementation of our distributed firewalling in wireless ad-hoc networks. The first design choice we made was to integrate the distributed firewall with the OLSR routing protocol. OLSR reflects the stateful and proactive nature of social networks, when a node enters an OLSR-based network it will incrementally build its routing table with routes to all the nodes in the network. If the OLSR packets are enriched with more information (the user-id, the current *state* ...) the user will be informed about the others in the network without the need to run additional discovery protocols.

OLSR also naturally supports efficient broadcast traffic and, with some modifications, multicast traffic [9] which permits typical social network applications, such as group chat.

We target a scenario where the nodes are carried by users, so that nodes are mainly moving with pedestrian speed. For the simulations, we implemented a mobility model that corresponds to human social behaviors [10] which are generally more predictable than random mobility models, but at the same time they are also more realistic.

Moreover, the typical usage pattern of social networks is to keep the social application as a background task of some other more important task for sessions that last from tens of minutes to hours, which allows to trade-off start-up time with other requirements. If the mobility of the nodes is low and some latency from initial connection to full connectivity is tolerable, richer information can be added to TC and HELLO messages trading the increased size with a lower repetition frequency as in fisheye strategies.

Lastly, to avoid attacks based on the distribution of fake rule-sets OLSR messages can be hardened with cryptography, using a centralized or a distributed model as described in various works in literature (see the works cited in [11]).

All these reasons made OLSR a suitable choice for the scenario and it also fits most other applications.

A. Evaluation Metrics

For the evaluation of the performance of our distributed firewall and rule-set reduction policies we use two metrics, metric M1 measures the false negatives on forward chain for drop targets and is meaningful to evaluate the efficiency of rule-set reduction policy applied by a node. In the simulation, every node keeps a cache of all the known rules and a separate list of enforced rules (reduced rule-set), M1 is measured per node as the fraction of packets that would match a rule in the full rule-set but do not match any in the reduced one.

Metric M2 meaningful to evaluate the efficiency of traffic limiting in mixed ad-hoc/mesh networks where only a subset of the nodes implement firewalling. It is calculated as the fraction of packets that reach their final destination IP but should have been filtered on their multi-hop path. Contrary to M1, it is calculated only at the final communication endpoints. In practice, M2 is calculated running a simulation batch without active firewalls and then comparing the results of the same runs with firewall actived on some nodes.

B. Preliminary results

Using a modified version of Inet module for Omnet++ simulator the integration between a network firewall and OLSR routing has been developed. Filters are matched against IP addresses and UDP/TCP ports, a pre-shared list of rule-sets is defined off-line and has an associated 8-bit numeric ID, each rule matches a UDP/TDP port.

Each node implements its own rule-set in the input chain for traffic destined to its own IP, and adds the corresponding ID to HELLO messages. MPR nodes in turn forward the ID together with the advertised address of the MPR selector in TC messages. Any node receiving a TC or HELLO can enforce the rule-sets referenced in the packet in its forward chain. Each rule-set is enforced matching in the IP packets the destination

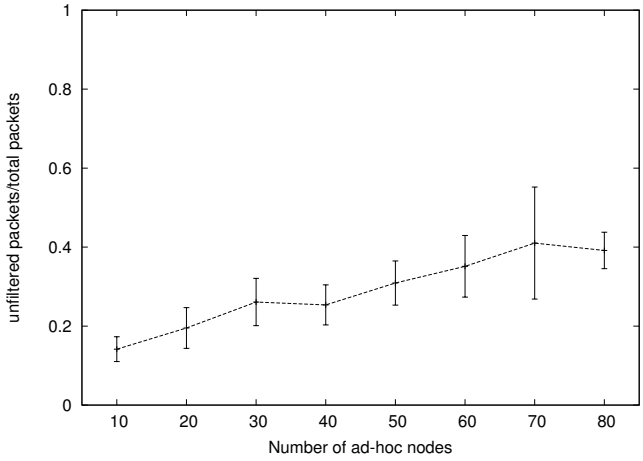

Figure 2. Metric M2: ratio between the number of packets filtered by mesh nodes when firewall is active and total number of packets received when firewall is not active

IP address that corresponds to the originator (for HELLO messages) or the advertised address (for TC messages). After a transitory phase, all the nodes in the network know the rules chosen by the other nodes and may enforce them all or not.

Two scenarios have been simulated so far, the first one is a mesh/ad-hoc scenario with 15 mesh nodes in a 5x3 grid, and a variable number of ad hoc nodes from 10 to 80. Each scenario is repeated with two configurations, in the first, the 15 mesh nodes enforce all the rule-sets (in the forward chain) while the ad-hoc nodes have no rules in the forward chain, in the second also the mesh nodes do not filter in the forward chain.

This scenario has been simulated in order to test metric M2 when there is an infrastructure network that tries to implement a specific rule-set and some ad-hoc nodes that generate and receive traffic. The simulated area is kept constant and the generated traffic is uniformly distributed among the possible ad-hoc nodes. The mesh nodes have a higher communication radius so that they are preferred for long paths, but increasing the density of ad hoc nodes makes it easier for a node to have routes that do not involve mesh nodes. For this reason Figure 2 shows that increasing the density, the number of packets that arrive at final destination increases (the firewall is more inefficient). We expect these results to be quite dependent on the scenario; however, even with only 10 different runs as in Figure 2 the standard deviation (reported in the plot as error bars) is low enough to observe that when mesh nodes enforce filters at least a 50% efficiency is reached, which is an encouraging result.

The second scenario illustrates the results of metric M1 applied to a possible strategy for the reduction of rule-set sizes and of the number of filtering nodes. This is accomplished with a tighter integration between OLSR routing and filtering and can be applied to purely ad-hoc networks. In this case, only the nodes that have a number of MPR selector above a given threshold actively filter and use the entire rule-set.

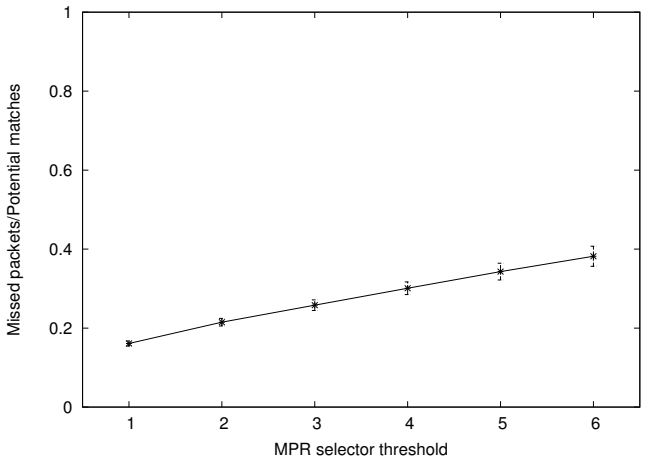

Figure 3. Metric M1: average efficiency of the rule-set reduction strategy

The other nodes use a reduced filter where all the rules that correspond to routes longer than two hops are removed. The intuition behind this strategy is that, under the assumption of uniformly distributed traffic, the more MPR selectors a node has the more traffic it will route. Metric M2 is zero, since no packet arrives at its final destination if there are rules to filter it. This is because traffic that passes through nodes with a MPR selector set size higher than the threshold will be fully filtered, but if a node has not enough MPR selectors it implements at least rules for 1 and 2 hops routes so it will not let traffic to be filtered reach the final destination. Metric M1 shows instead the efficiency of the global filtering function. Figure 3 shows that in a network composed of 55 ad-hoc nodes this strategy is able to filter on average more than 60% of the packets. Note that this metric is calculated per node per hop, not at final destination like M2. When the MPR selector threshold is set to one, all the MPRs apply the full rule-set, so that unfiltered packets are concentrated on the first two hops of the traffic. With a higher threshold, less nodes apply the full rule-set but it can be seen that the decay of performance is no worse than linear. The number of MPR selector is a very simple heuristic, since OLSR has a proactive knowledge of the topology more complex ones can be used.

Obviously there is an uneven distribution of filtering load, which consequently makes CPU load uneven between nodes. Still, having less nodes with full rule-set has the consequences of reducing the time needed for the routing decision and thus limit the average latency on a path (the evaluation of a large rule-set may need several milliseconds on a low-performance device).

IV. Conclusion

The use of ad-hoc networks to provide services on-demand, i.e., when and where they are needed, is slowly rolling out and changing the Internet scenario.

This changed scenario expose nodes and users to novel threats and problems; for instance all the "security and privacy" guaranteed by traditional NAT-plus-firewall architectures may not be applicable at all.

In this work, we have first introduced problems related to distributed firewalling in ad-hoc networks, taking as reference sample application a social network.

A prototype implementation over OLSR in Omnet++ has been developed and initial results for some metrics of interest show that this architecture is viable and sustainable, though many issues remains open.

So far we have focused on the where and how to apply filtering rules in a cooperative scenario; however, problems related to rules generation, their distribution and finally their enforcement in non-cooperative environments remain open and draw a clear road-map for future research.

Acknowledgements

This work is funded by The Trentino programme of research, training and mobility of post-doctoral researchers, incoming Post-docs 2010, CALL 1, PCOFUND-GA-2008-226070. For more information and the source code of the simulator visit http://pervacy.eu.

References

[1] D. E. Taylor, "Survey and taxonomy of packet classification techniques," *ACM Comput. Surv.*, vol. 37, pp. 238–275, September 2005.

[2] P. Ayuso, "Netfilters connection tracking system," *LOGIN: The USENIX magazine*, vol. 31, pp. 34–39, June 2006.

[3] S. Ioannidis, A. D. Keromytis, S. M. Bellovin, and J. M. Smith, "Implementing a distributed firewall," in *ACM Conference on Computer and Communications Security*, Athens, Greece, November 2000.

[4] R. Fantacci, L. Maccari, P. Ayuso, and R. Gasca, "Efficient packet filtering in wireless ad hoc networks," *Communications Magazine, IEEE*, vol. 46, no. 2, pp. 104 –110, February 2008.

[5] M. Taghizadeh, A. Khakpour, A. Liu, and S. Biswas, "Collaborative firewalling in wireless networks," in *IEEE INFOCOM*, Shanghai, China, April 2011.

[6] H. Zhao, C.-K. Chau, and S. M. Bellovin, "ROFL: Routing as the firewall layer," in *New Security Paradigms Workshop*, Lake Tahoe, USA, September 2008.

[7] H. Zhang, B. DeCleene, J. Kurose, and D. Towsley, "Bootstrapping deny-by-default access control for mobile ad-hoc networks," *Military Communications Conference, IEEE MILCOM*, San Diego, USA, November 2008

[8] M. Alicherry, A. D. Keromytis, and A. Stavrou, "Evaluating a collaborative defense architecture for manets," in *IEEE international conference on Internet multimedia services architecture and applications, IMSAA'09*. Bancalore, India, December 2009.

[9] A. Laouiti, P. Jacquet, P. Minet, L. Viennot, T. Clausen, and C. Adjih, "Multicast Optimized Link State Routing," INRIA, Research Report RR-4721, 2003. [Online]. Available: http://hal.inria.fr/inria-00071865/en/

[10] M. Musolesi and C. Mascolo, "A community based mobility model for ad hoc network research," in *International workshop on Multi-hop ad hoc networks: from theory to reality, REALMAN*. Florence, Italy, May 2006.

[11] E. A. Panaousis, G. Drew, G. P. Millar, T. A. Ramrekha, and C. Politis, "A testbed implementation for securing olsr in mobile ad hoc networks," *International Journal of Network Security & Its Applications*, Vol. 2, No. 4, 2010.