On the Use of Admission Control for Better Quality of Security

Svetlana Radosavac, Ulaş C. Kozat, and James Kempf
DoCoMo Communications Laboratories USA, Inc.
Palo Alto, CA 94304
Email: {sradosavac,kozat,kempf}@docomolabs-usa.com

***Abstract*—We propose an admission control policy that admits users into a public access network as soon as possible while limiting the overall security impact on the network and other users. In our model, each user has a particular reputation level when first requesting network access. Before admitting a user into the network, the initial risk of a user is assessed by the admission control system using past history and a scanning of the user's device which delays the user's admission into the network and updates the user's reputation level accordingly. We formulate the trade-off between the admission delay and security risk as a convex optimization problem, which can be solved for an admission control policy. The evaluation suggests that our approach can substantially increase the system security for public access networks while minimizing admission delay, in contrast to current approaches widely used in enterprise networks. The proposed framework extends the traditional quality of service-based admission control mechanisms with a well-defined notion of quality of security.**

I. Introduction

A secure public access network, in which user access and traffic is more tightly controlled than in today's public Internet, could have much value in improving the security of transactions between users and security-sensitive businesses such as banks. Creating a secure public accessible network is still an open research problem, however. The problem is especially acute for mobile users who frequently switch between different networks. Every time a user leaves and returns to a secure network, new security vulnerabilities potentially tag along. In enterprise networks, the traditional approach is to isolate a returning user on a separate VLAN segment and perform a full system scan, patching any discovered vulnerabilities, before the user's device is admitted [1]. However, this approach can introduce excessive delays every time a user wants to access the network. In mobile public access networks, where users often connect for short transactions and frequently switch between different networks, for example WLAN and 3G, long admission delays for full systems scans are unacceptable. In this paper, we describe an admission control policy that better trades off admission delay and system security.

The particular security concern in our work is containing the risk of distributed denial of service (DDoS) attacks. DDoS attacks are launched by thousands of compromised user machines (botnets) that flood ISP networks with traffic, effectively rendering the networks or services unreachable. In Feburary 2000 [5], a DDoS attack on major Internet services alerted the public to the problem. The Symantec Internet Security Threat Report [6] estimates 6 million infected computers worldwide are connected to botnets.

The basic premise of our admission control design is that, unlike current Internet service, the network operator of a secure network service is responsible for compensating users who become victims of DDoS attacks. In turn, users are willing to pay a premium above the tariff they would pay for access to flat-rate Internet service, and are willing to undergo a variety of security measures to ensure that they don't pose an undue risk to the network. The network operator consequently must bear risk of DDoS attacks either by absorbing the risk or transferring it to a third party using insurance. From the network operator point of view, it is critical to admit users such that the expected damage is below the damage threshold that can be tolerated by the network operator. While the measures that we assume to be necessary to maintain a secure public access network may seem somewhat intrusive, it is important to emphasize that without sufficient knowledge about the risks posed by entering users, network operators cannot guarantee the security of the network.

For quantifying the likelihood of a botnet, the proposed admission control policy assesses the probability of a user becoming part of a botnet. This probability is computed according to the reputation of the user, which is derived from the past behavior of the user as well as from real-time scanning and device inspection before and after the user is admitted into the network. We assume that the user's reputation improves as a function of the scanning time and that the utility of a user decreases as its admission delay increases. Accordingly, we pose the admission control policy as a solution to a constrained convex optimization problem, where the objective is to maximize the sum utility of all admitted users. The system constraints are defined with respect to the damage threshold on the expected damage of admitted users and the progress of the scanning of new users. Whenever the expected damage is sufficiently low, a new user can be admitted into the network.

The expected damage becomes sufficiently low as: (i) the reputation of admitted users increases through monitoring of their traffic for DDoS traffic and through additional security measurers the users take; (ii) users with lower reputations depart; (iii) the capacity increases inside the network and at the edges allow higher levels of DDoS traffic to be accommodated without user impact and (iv) the operator can accommodate the increased financial damage of a DDoS attack through

978-1-4244-3435-0/09/$25.00 ©2009 IEEE

accumulation of enough wealth to compensate victims or through transferring the risk by purchasing insurance, leading to a higher operator damage threshold.

We provide a general solution to the optimization problem over arbitrary concave utility functions and present a closed form solution based on a particular analytical form of the utility function, and an efficient admission control algorithm. The solution to the optimization problem states which user is to be admitted and when to the network given a particular set of already admitted users. The admission control policy updates its admission time decisions after each new user arrival. Our evaluation indicates that we can substantially improve the admission delay and substantially increase the sum utility over static approaches that perform a complete scan before admitting a new user.

The rest of the paper is organized as follows. In Section II, we discuss the related work on risk modeling of DDoS attacks and security interdependence of network users. We provide the details of our system model in Section III and present the optimization problem and its solution as an admission control algorithm in Section IV. In Section V, we present performance evaluation results of our admission mechanism and conclude the paper in Section VI.

II. Related Work

The problem of the optimal attack on a network was studied under different conditions in [2]. Cunningham considers the problem of an optimal DDoS attack as a problem of minimizing the ratio of the edge destruction cost to the number of disconnected components, and defines the *strength* of a network as a measure of the resistance of the network to such attacks. Efficient algorithms are provided for the optimal attack problem, the problem of computing the strength, and the problem of finding a minimum cost reinforcement to achieve a desired strength. Following that, the attacker's goal is solved by separating vertices from a fixed central vertex.

Most Network Admission Control (NAC) products currently available, such as Cisco NAC Appliance system [1], perform checks of device security (checking whether the OS is patched, if anti-virus software exists on the inspected machine and if the anti-virus software is running), identity (checking whether the user is authorized to use the network and if there are any restrictions on the user's access of available resources) and network security (whether the required security policies are established within the network, whether non-compliant devices are quarantined, and if it is necessary to patch network devices). After such actions are performed, the device undergoes another security check and if all the necessary fixes are applied, the device is admitted into the network.

More recently there is a growing body of literature that tries to model and address the network security problem using risk modeling, network economics and incentives. Along these lines compensation-based methods such as cyber insurance and incentive mechanisms for improving network security have been proposed. However, very few businesses take advantage of such policies. As a response to an alarmingly low adoption rate of cyber insurance, [3] provides a detailed description of problems that arise in this field. Our work is relevant to healthy deployment of cyber-insurance, since the network operator explicitly constrains the security risks of admitted user and could potentially use cyber insurance to compensate victims of DDoS attack.

III. System Model

We assume the system consists of K users that have already been admitted into the network and N users that are waiting for admission. Whenever a user is disconnected from the network due to an attack, it suffers financial losses and the network operator compensates the user for the incurred damages. On the other hand, the utility of the network increases as more users are admitted. Therefore, the goal of the network is to admit as many users as quickly as possible while keeping expected damages below a tolerable threshold. To achieve that, the network needs to construct an admission policy that admits users based on a security assessment and the security risk they impose to the network.

Each user $u_i, i \in \{1, \ldots, N\}$ that attempts to join the secure network is characterized with two parameters: reputation p_i, $p_i \in [0, p_{i,max}]$ and traffic injection rate r_i (i.e. the user is allowed to inject traffic no faster than this rate). In this work we assume $p_{i,max} = 1$. Reputation p_i signifies the trust level put on that user by the network. When user u_i requests access to the secure network, the system determines its initial reputation value $p_{i,0}$ based on the results of real-time scanning the user device, past interactions between the user and the network, etc. After $p_{i,0}$ is determined for all users waiting for admission, the admission control policy mechanism makes an admission decision for each user by computing the overall risk after the admission for that particular user. The reputation of an admitted user is assessed and updated in real-time starting from an initial value at the time of the user's arrival:

$$p_i = p_{i0} + g(\tau_i), i = 1, \ldots, N \tag{1}$$

where $g(\tau_i)$ is a non-negative non-decreasing function of admission delay τ_i, i.e., as the access delay increases the system either discovers that the specific user is secure and the reputation increases or discovers that the user does not have the required properties and forces the user to update the device, which results in increase of the user's reputation (otherwise the reputation remains the same).

We capture the security threat of a given subset $\mathbf{B_i} = \{u_{i1}, \ldots, u_{im}\}$ of users $u_{i1}, \ldots, u_{im}$ (where the subset $\mathbf{B_i}$ is created from the set of already admitted users and users that are waiting for admission) by the sum rate $\sum_i = \sum_{j=1}^{m} r_{ij}$, where r_{ij} is the traffic injection rate of user u_{ij}. The damage $D(B_i)$ that can be caused by subset B_i is then modeled as a non-decreasing monotonic function of $\sum_i$, i.e. $D(B_i) = f(\sum_i)$. If we assume that the probability of a user becoming part of a botnet is independent of other users, the damage probability π_i of a particular subset of users $\mathbf{B_i}$ can be computed as:

$$\pi_i = \prod_{j \in B_i} (1 - p_j) \ . \tag{2}$$

Here, $(1-p_j)$ is used to measure the likelihood/probability of user j becoming a member of a malicious subset B_i.

The admission control policy guarantees that expected damage over all possible subsets of admitted users is less then a threshold Γ_{th}, which is computed using the a-priori assessment of the damage cost that is tolerable to the network operator and other users. The admission control policy mechanism computes the expected damage as:

$$E_B[D] = \sum_{i\in B} \pi_i D(B_i) \tag{3}$$

where the expectation is taken over the set $\mathbf{B}$ of all possible subsets $\mathbf{B}_i$.

Each user i is assumed to have a concave and decreasing utility function $U_i(\tau_i)$ in τ_i. Such a utility function definition is reasonable because (i) the less time a user waits to be admitted into the system, the more satisfied the user is and (ii) user's utility decreases faster at larger admission delay points with no positive utility above a maximum tolerable admission delay.

IV. Optimization Problem and Admission Control Policy

The admission control policy mechanism decides whether to admit a user at a particular time by solving the following constraint optimization problem:

$$max \textstyle\sum_i^N U_i(\tau_i) \tag{4}$$

$$\text{s.t.}$$

$$E_B[D] \leq \Gamma_{th} \tag{5}$$

$$p_i = p_{i0} + g(\tau_i), i = 1,\ldots,N \tag{6}$$

$$0 \leq p_i \leq 1, i = 1,\ldots,N \tag{7}$$

In other words, the objective of the admission mechanism is to maximize the sum utility of users admitted into the system, where the individual utilities depend on the admission delay of that particular user.

The constraint (5) reflects the cost of expected damage to the network over the set of possible botnets. $g(\tau_i)$ in (6) (and consequently p_i) is a non-negative non-decreasing function of admission delay τ_i. When a user requests to connect to the network, its initial reputation value is securely obtained by an off-line evaluation and the admission control policy mechanism starts scanning the user's device. When the device passes the scan, the user's reputation increases. Otherwise, the device is quarantined, the necessary security patches are installed and the scanning process is repeated.

In the rest of the paper we assume that $g(\tau_i) = \alpha\tau_i$, where α is a positive constant for the sake of simplicity. Note that in this optimization framework, users do not need to wait until a full scanning is completed and/or necessary patches are applied. If the risk they impose on the network as captured by Eq. (5) is acceptable and there is greater gain in admitting the user earlier rather than performing a full scan in accordance with the utility maximization, then the user can be admitted earlier. However, the scanning process and reputation update continues even after the user is admitted to the system.

To further simplify the above optimization problem, we assume the following form of $D(B_i)$: $D(B_i) = s \times \sum_i$, i.e. the damage $D(B_i)$ is a linear function of sum rate of B_i. Without loss of generality, we set $s = 1$ since we can normalize the constraint equation by s. For determining the value of $E_B[D]$ in Eq. (3) we use an upper bound that is convex in p_i. One such function is of the form

$$E_B[D] = \gamma \sum_{i=1}^{N} (1-p_i) r_i \ , \tag{8}$$

where γ is a real value between 1 and 2^N. Note that $\gamma = 2^N$ is a trivial upper-bound, but in general it is a loose one especially for large N. Smaller γ values can be used to more tighly upper-bound the expected damage. Here r_i denotes the allowed connection rate of the i^{th} user.

With the linear approximations for the constraints, the optimization problem becomes a convex optimization problem. Accordingly, we can define the following Lagrangian function:

$$\phi(\lambda,\mu) = \sum_{i=1}^{N} U_i(\tau_i) - \lambda\cdot(\gamma\sum_{i=1}^{N}(1-p_{i0}-\alpha\tau_i)r_i - \Gamma_{th}) - \sum_{i=1}^{N} \mu_i \cdot (p_{i0} + \alpha\tau_i - 1) \ ,$$

where λ and μ_i are the Lagrange multipliers. Solving for the Kuhn-Tucker conditions reveals that:

$$\sum_{i\in A} U_i^{'-1}[-\alpha\gamma r_i \lambda] r_i = \frac{1}{\alpha}\left(\sum_{i\in A}(1-p_{i0})r_i - \frac{\Gamma_{th}}{\gamma}\right) \tag{9}$$

where $U_i'(.)$ is the derivative of U_i and $U_i^{'-1}(.)$ is the inverse function of $U_i'(.)$.

Let A be the set of users who are admitted into the system before they reach the maximum reputation level. Once λ is computed, the admission delay of each user i can be computed as:

$$\tau_i = \begin{cases} \tau_{i,max}; \ i \notin A \\ U_i^{'-1}[-\alpha\gamma r_i \lambda]; \ i \in A \end{cases} \tag{10}$$

Here, $\tau_{i,max} = (1-p_{i0})/\alpha$ represents the maximum delay user i can observe before being admitted into the system (after $\tau_{i,max}$ the user's reputation becomes one, thus it does not pose a security risk).

We still need to determine the value of the set A to find the admission times for each user. First, observe that for each $i \in A$, we have the inequality

$$\tau_i = U_i^{'-1}[-\alpha\gamma r_i \lambda] < \tau_{i,max} \ ,$$

or equivalently,

$$\lambda < \frac{-U_i^{'}(\tau_{i,max})}{\alpha\gamma r_i} = \lambda_i \ , \tag{11}$$

due to the fact that $U_i^{'}(x)$ is a decreasing function of x. We refer to λ_i as the user priority function. Note that the lower λ values would result in admitting more users into the system since more user would satisfy the above inequality. A lower λ would also decrease the admission delay of each user since

$U_i'(x)$ is a decreasing function of x and so $U_i^{'-1}(-x)$ is an increasing function of x. Therefore, a lower λ value implies a lower admission delay for users in set A. As a result, the sum utility increases. The optimum solution finds the smallest $\lambda(A) > 0$ such that the above constraints are satisfied.

Note that if a user with λ_i is in A, then all j such that $\lambda_j \geq \lambda_i$ must be in A. Let us sort all users in increasing λ_i such that $\lambda_{j_1} \leq \lambda_{j_2} \leq \ldots \leq \lambda_{j_N}$. Set A can then be one of the following $(N+1)$ subsets $A(0) = \{j_1,\ldots,j_N\}$, $A(1) = \{j_2,\ldots,j_N\}$, $A(2) = \{j_3,\ldots,j_N\}$, $\ldots$, $A(N-1) = \{j_N\}$, $A(N) = \emptyset$. The admission control policy executes the following algorithm:

```
for m=0 to N
        $A := A(m)$;
        Compute $\lambda$ according to Eq. (9);
        Set $A' = \{\forall i : \lambda_i > \lambda\}$;
        If $A' \equiv A$ set $\lambda^* = \lambda$ and exit;
        Else continue to next iteration;
end
```

After the algorithm halts, the admission time τ_i of each user i is computed according to Eq. (10) using $\lambda = \lambda^*$ and last A. The users are then admitted at time $\tau_{arr,i} + \tau_i$, where $\tau_{arr,i}$ is the arrival time of user i into the system.

In a typical case, where there are multiple arrivals of users over a certain period of time, each new user arrival triggers the re-execution of the algorithm for the users who are waiting for admission either because their admission time is not reached or because their admission time is not yet computed (e.g., the newly arrived users). Suppose Z denotes the set of users at time t who are already admitted into the network, not yet departed, and have reputation less than one, i.e., $Z = \{\forall i : \tau_{arr,i} + \tau_i < t < \tau_{depart,i} \wedge p_i < 1\}$. Also, let Ψ be the set of all possible subsets of Z. The first constraint given by (5) can then be rewritten as $E_B[D] \leq \Gamma_{th} - E_\Psi[D]$. In other words, the damage threshold is reduced by the expected damage that can be caused by different subsets of users who are already admitted into the network. We can compute this damage as

$$E_\Psi[D] = \gamma \sum_{i \in Z} (1 - p_i) r_i. \tag{12}$$

After the new conditions are taken into account - i.e., the expected damage is updated with the newly arrived users, their reputations, and traffic injection rates - the admission times for the users who are still waiting to enter the network are computed/recomputed using the same algorithm as before.

A. Solution for a specific utility function

In order to provide further insight about the proposed admission mechanism and its performance, in this section we provide a closed form solution of the optimization problem for a specific utility function $U_i(\cdot)$. Since we assume that the user's utility function is a concave decreasing function of τ_i, we use the following utility function:

$$U_i(\tau_i) = \beta\Big(1 + ln\Big(1 - \frac{\tau_i}{\tau_{max}}\Big)\Big) \tag{13}$$

Using this expression for the utility, we can now calculate the exact values of priority λ_i from Eq. (11):

$$\lambda_i = \frac{\beta}{\alpha\gamma r_i(\tau_{max} - \tau_{i,max})} \tag{14}$$

where τ_{max} represents the maximum tolerable access delay.

For this utility function, we observe that the users with a higher upload rate have lower λ_i values. Hence, they are more likely to be excluded from set A and wait until the end of a full scan. We also see that users with lower $\tau_{i,max}$ (higher initial reputation) are also likely to be delayed until their systems are fully scanned. This might seem surprising, since the users with a higher initial reputation value are likely to pose less security risk. However, the key is the utility function: when users have high initial reputation value, their maximum delays $\tau_{i,max}$ are already small and their utilities are little impacted by the extra delay. Hence, from the network point of view, it is reasonable to preferentially opt for eliminating the security risks of lower reputation users with full scanning and patching.

Following the steps of the admission algorithm outlined in the previous section and utilizing the expression for λ_i from Eq. 14, we now compute the expression for λ from Eq. 9:

$$\lambda(A(m)) = \frac{\beta}{\frac{1}{|A(m)|}\sum_{i \in A(m)} \frac{\beta}{\lambda_i} + \frac{\Gamma_{th}}{|A(m)|}}. \tag{15}$$

and compute the corresponding $\lambda^\star$, as it was outlined in the previous section.

Since the utility function is of form Eq. 13, we have:

$$U_i^{'-1}[-\alpha\gamma r_i \lambda^\star] = \tau_{max} - \frac{\beta}{\alpha\gamma r_i \lambda^\star}. \tag{16}$$

and the access delay is:

$$\tau_i = \begin{cases} \tau_{i,max}; \ i \notin A \\ \tau_{max} - \frac{\beta}{\alpha\gamma r_i \lambda^\star}; \ i \in A \end{cases} \tag{17}$$

In the following section we present an experimental evaluation of the system performance using simulation when the utility function from Eq. (13) is assumed for the arriving users.

V. Simulation Results

We conducted simulations to evaluate the proposed dynamic admission control mechanism. Along with our dynamic admission control mechanism, we also simulated the static admission control mechanism used in enterprise networks, where each user is fully scanned until the user's reputation reaches the maximum value $p_{max} = 1$. The user arrival process is Poisson with rate $\lambda = 10 \ \frac{user}{second}$. We assume that the user's initial reputation p_0 is uniform in the interval $[0,1]$. Each user's traffic injection rate is randomly chosen from the interval $[100,1000]$ and the lifetime of a user after admission into the network follows an exponential process with mean 100s.

The static admission mechanism increases the user's reputation linearly in time, i.e. after the scanning process is completed the network has a precise estimate of the user's reputation. Such mechanism does not allow users to access the network before the scanning process is completed and the user also must install all the necessary patches and updates prior to admission.

The results for the static admission mechanism are presented in Fig. 1. In the simulation, we assume that the reputation of a user increases linearly from p_0 to $p_{max} = 1$ with parameter $\alpha = 0.2$, which results in the maximal static admission delay of $5s$. As expected, the delay distribution is almost uniform between minimum and maximum delays.

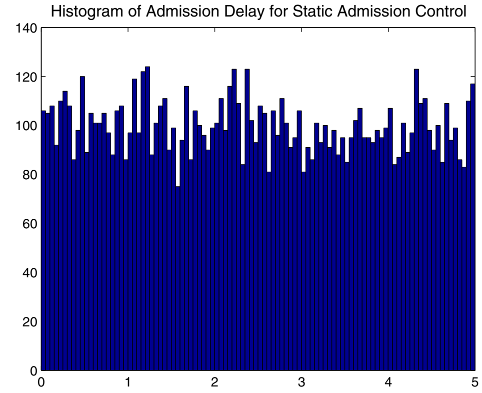

Fig. 1. Static Admission Control.

Fig. 2 represents the admission delay for our proposed admission mechanism. Due to the ability of the network to tolerate some damage, a significant number of users are admitted with noticeably lower delay. The network can admit users that have reputation $p < p_{max}$ if the overall assessed damage is lower than the specified threshold. Furthermore, the admission threshold value is updated as the user exits the network, so if a user with a high traffic injection rate leaves the network or reaches p_{max}, a new set of users can be admitted.

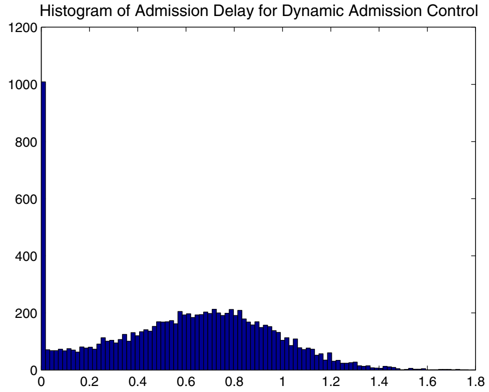

Fig. 2. Dynamic Admission Control.

In order to obtain further intuition from our results, we also compare the mean access delay for both mechanisms by varying the initial reputation distribution. The results of the comparison are presented in Fig. 3. The initial value of reputation is now chosen from the interval $[0, i]$, $i \in [0.1, 1]$. We observe that the admission delay for both schemes decreases as the interval over which p_0 is chosen increases. In the case of static admission control, a certain portion of users who are assigned higher initial reputation values wait for a shorter period of time for admission since p_{max} increases. In the case of dynamic admission control it is now possible to fit the users having high p_0 together with the users having low p_0, resulting in lower access delays. We observe from Fig. 3 that the percentage decrease in access delay with the increase of the initial reputation interval is slightly higher in static admission mechanism (47%) than for the dynamic mechanism (41%), which is due to the fact that the admission delay for the static admission mechanism depends only on the initial reputation.

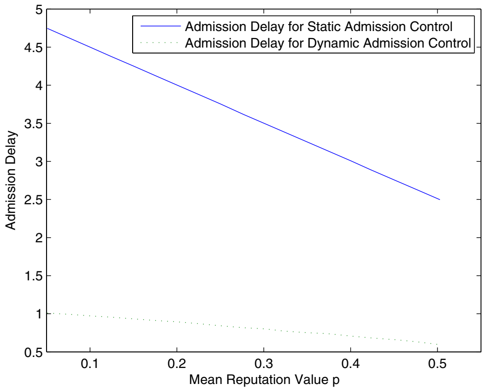

Fig. 3. Admission Delay for various values if initial reputation.

VI. Conclusion

In this work, we proposed an admission control policy for securing public access networks. We posed the admission control question as a constrained utility maximization problem. Our solution framework constrains the overall security risks of the admitted users, and within that constraint aims to maximally improve the admission delay into the network. The admission control policy proposed in this work not only determines whether a user is admissible or not, but also when the user is admissible. We provided an efficient algorithm for admission control under very generic assumptions. We also presented further results with specific utility function definitions. Our overall evaluation suggests that indeed the proposed admission control policy can improve both the mean delay as well as the delay distribution of users over more traditional and static admission control strategies that are agnostic to the network operator's ability to compensate users for damage, as long as the risks are properly managed.

References

[1] Cisco NAC Appliance Data Sheet. http://www.cisco.com/en/US/prod/collateral/vpndevc/ps5707/ps8418/ps6128/product_data_sheet0900aecd806e98c9.html

[2] W. H. Cunningham, *Optimal attack and reinforcement of a network*, J. ACM, Vol.32, No. 3, 1985.

[3] W. S. Baer and A. Parkinson, *Cyberinsurance in IT Security Management*, p.p. 50-56, IEEE Security and Privacy, Vol. 5, No. 3, May 2007

[4] B. Hillier *The Economics of Asymmetric Insurance*, Palgrave Macmillan, 1997.

[5] K. T. Tran and R. L. Rundle. Hackers attack major web sites, shutting amazon, buy.com, ebay, Feb. 2000.

[6] Symantec internet security threat report, Sept. 2007.