# Worst Case Attacks on Distributed Resources Systems

Hanoch Levy
Blavatnik School of Computer Science
Tel-Aviv University
hanoch@tauex.tau.ac.il

Jhonatan Tavori
Blavatnik School of Computer Science
Tel-Aviv University
jhonatant@mail.tau.ac.il

## 1. INTRODUCTION

How should an attacker, who wishes to hurt (deny) service, attack resources on a geographically distributed system in order to maximize the damage inflicted? Should attack efforts focus on a small number of regions (sites) or rather spread over many regions to yield wide spread effect? Consequently, how resources should be optimized as to minimize the gain of a damage-maximizing attacker? These are the questions that motivate this research and are at its core.

The questions apply to a large variety of highly sensitive applications, both in the cyber domain and in the physical domain, such as management of resources on distributed clouds, allocation of base-station resources in cellular networks and control of electricity production plants on the electrical grid. For the sake of simplicity our presentation will focus on distributed cloud servers.

The answer to these questions is important for designing distributed resources systems properly. Placing the resources at such systems, in preparing for attacks, must be done in order to minimize the attack damage and under the assumption that the attacker aims at maximizing the damage. It is common knowledge that cyber attacks have become of major concern over the years, especially when the world has been shifting towards cloud solutions.

It is important to note that the question we address is how to maximize the damage inflicted on the *users* served by the resources, and *not* how to maximize the damage inflicted on the servers/resources, or at failing the system[1]. The fact that the user demand to these resources is *stochastic* is what makes the question challenging.

A number of studies dealt with resource allocation problems in the cloud (e.g., [1], [2] and [3]). [3] dealt with deriving optimal resource allocation to cope with faulty resources, resulting from either malfunction or from attacks. However, [3] inherently assumes that failures are random and are *not* tailored to maximize damage. [3] also provides the motivation for our work which aims at understanding how to optimize resource allocation in order to cope with *damage maximizing* attackers, and whether such optimal allocations inherently differ from those derived in [3].

The results reported in this work form a preliminary set of results that characterize the methodology of an optimal attack on distributed resource systems. We use a model where the user demand for resources/services is of arbitrary multi-dimensional distribution. We start the analysis by addressing deterministic attacks, namely where each attack agent inflicts deterministic damage (e.g. takes down a server). We then move to stochastic attacks, which are more realistic, where this damage is probabilistic (succeeds with probability $p$). We further consider a more sophisticated defence whereby the system may mitigate the attack damage by using a real-time shift of demands from one region to another. Our results show that for all these configurations, it is optimal for the attacker to focus efforts on a small number of regions.

Further work is described in Section 5.

## 2. THE MODEL

The system consists of $n$ regions numbered $1, 2, \ldots, n$. Let $L_i$ be the number of resources placed (servers) in region $i$, for $i \in \{1, \ldots, n\}$. The placed resources vector in the system is called an *allocation*, and is denoted by $\mathbf{L} = (L_1, \ldots, L_n)$.

Let $D_i$ be a discrete random variable denoting the *demand* (number of resources requests) in region $i$. We do not make any assumption on the distribution of $D_i$, that is, it can be of an arbitrary distribution, besides the trivial requirement that the support should be entirely non-negative. Further, we do not assume independence between the demands, namely $D_i$ and $D_j$ are not necessarily mutually independent. Finally, let the system's demand be denoted by $\mathbf{D} = (D_1, \ldots, D_n)$.

An attacker who wishes to hurt the the system will invest *attacking agents* in order to do so. Let $x_i$ by the number of attacking agents invested in region $i$ (such as viruses for example). The attack vector (or simply *attack*) on the network is $\mathbf{X} = (x_1, \ldots, x_n)$. $\|\mathbf{X}\| = \sum x_i$ is the *attack size*.

As $L_i$ is the number of allocated resources in the region $i$, we denote by $S(L_i, x_i)$ the *supply* of the resources in that region. $S(L_i, x_i)$ is a random variable and is the number of surviving resources after a possible failure due to an attack of size $x_i$ on a placement of $L_i$ resources. We note that the supply volume cannot be larger than the placement and cannot be negative. Hence, the support of the supply distribution $S(L_i, x_i)$ is $\{0, 1, \ldots, L_i\}$, for any $x_i$. For brevity, from now on $L_i$ and $S(L_i, x_i)$ will be referred to as the placement and supply under the attack $\mathbf{X}$ in region $i$, respectively.

[1]When one wishes to "fail a system" it is quite intuitive that the optimal strategy is to concentrate all efforts on the weakest defence point, such as militaries used to do when they aimed at "breaking a wall" to conquer a city. However, since we aim at maximizing the damage to the stochastic demand served by the system, optimal strategy is not obvious

## 2.1 Objective Function

Consider a request made in region $i$ under the supply $S(L_i, x_i)$. If the request is assigned to a resource in the region, then the request is called *satisfied*. We assume that the system is rewarded from assigning a request to a resource (and satisfying it). The reward for each such request in region $i$ is denoted $r_i$.

Let $d_i$ and $s(L_i, x_i)$ be realizations of the demand $D_i$ and the supply $S(L_i, x_i)$ in region $i$. The total reward derived in region $i$ under this realization is equal to $r_i \cdot \min\{s(L_i, x_i), d_i\}$.

Finally, the objective function of the attacker is to minimize the expected revenue of the system, $R(\mathbf{X})$, parameterized by the attack $\mathbf{X}$ where its volume is limited by $\|\mathbf{X}\| \leq x$:

$$\min_{\mathbf{X}:\|\mathbf{X}\|\leq x} R(\mathbf{X}) = \min_{\mathbf{X}:\|\mathbf{X}\|\leq x} \sum_{i=1}^{n} r_i \cdot \mathbb{E}[\min\{S(L_i, x_i), D_i\}]. \quad (1)$$

# 3. OPTIMAL STRATEGY

In this section we prove that concentrating the attack is an optimal strategy for the attacker. We start by looking at deterministic attacks. Afterwards, we move to a stochastic attack model with a binomial supply (of concurrent attack agents).

## 3.1 Deterministic Attacks

In a deterministic attack each attacking agent disables a deterministic (and fixed) number of system resources. For simplicity of presentation we may assume that this number is 1. Namely, given that $L_i$ resources were placed in region $i$, the number of surviving resources is $S(L_i, x_i) = L_i - x_i$. Using the fact that for any non-negative random variable $Z$ and any positive integer $z$ the following holds: $\mathbb{E}[\min\{Z, z\}] = \sum_{i=1}^{z} \Pr[Z \geq i]$, we get that the expected revenue of the whole system, as a result of the attack $\mathbf{X}$, is:

$$R(\mathbf{X}) = \sum_{i=1}^{n} \mathbb{E}[\min\{S(L_i, x_i), D_i\}] = \sum_{i=1}^{n} \sum_{j=1}^{L_i - x_i} \Pr[D_i \geq j]. \quad (2)$$

Let $\Delta_i(x_i)$ be the *marginal damage* due to adding the $x_i$th attacking agent in region $i$ (comparing to only launching $x_i - 1$ agents). Namely, $\Delta_i(x_i) = -(R(\mathbf{X}) - R(\mathbf{X}^{i-1}))$, where $\mathbf{X}^{i-1} = (x_1, \ldots, x_i - 1, \ldots, x_n)$. For the deterministic attack, we have:

$$\Delta_i(x_i) = r_i \cdot \Pr[D_i \geq L_i - x_i + 1]. \quad (3)$$

Note that: $R(\mathbf{X}) = R(\bar{\mathbf{0}}) - \sum_{i=1}^{n} \sum_{j=1}^{x_i} \Delta_i(j)$, where $\bar{\mathbf{0}}$ is the "empty attack" (i.e., no attacking agents in use).

*Definition 1.* An attack realization $\mathbf{X} = (x_1, \ldots, x_n)$ is a *concentrated attack* if for one region at most (or none), say $j$, it holds that $0 < x_j < L_j$. For all other regions it holds that $x_i = 0$ or $x_i = L_i$.

We move to prove the main result of this section, stating that concentrating the attack is the optimal strategy for a deterministic attack. In our proof we use the fact that for any $i, j$: $\Delta_i(j) \leq \Delta_i(j+1)$, resulting from (3) and the monotonicity of cumulative distribution functions.

*Theorem 1.* There exists a concentrated-attack vector $\hat{\mathbf{X}}$, $\|\hat{\mathbf{X}}\| \leq x$, for which:

$$R(\hat{\mathbf{X}}) = \min_{\mathbf{X}:\|\mathbf{X}\|\leq x} R(\mathbf{X}) \quad (4)$$

PROOF. Let $\mathbf{X'} = (x_1, x_2, \ldots, x_n)$ be an attack vector such that $R(\mathbf{X'}) = \min_{\mathbf{X}:\|\mathbf{X}\|\leq x} R(\mathbf{X})$. Assume that there exist two regions, $i_1, i_2$ such that:

$$0 < x_{i_1} < L_{i_1}, 0 < x_{i_2} < L_{i_2}$$

(otherwise, the proof is complete). If $\Delta_{i_1}(x_{i_1}) \neq \Delta_{i_2}(x_{i_2})$, then w.l.o.g $\Delta_{i_1}(x_{i_1}) < \Delta_{i_2}(x_{i_2})$. By moving an attacking agent from region $i_1$ to region $i_2$, we get an attack with higher damage since by the monotonicity of the marginal damage, $\Delta_{i_1}(x_{i_1}) < \Delta_{i_2}(x_{i_2}) \leq \Delta_{i_2}(x_{i_2} + 1)$ – a contradiction to optimality of $\mathbf{X'}$. If, on the other hand, $\Delta_{i_1}(x_{i_l}) = \Delta_{i_2}(x_{i_2})$ we pick w.l.o.g region $i_1$ and move attacking agents from it to $i_2$, until we either reach $\Delta_{i_1}(x_{i_l}) < \Delta_{i_2}(x_{i_2})$, or we reach $x_{i_1} = 0$ or $x_{i_2} = L_{i_2}$. In the first case, as before, we have a higher damage to the system – a contradiction to optimality of $\mathbf{X'}$. In the latter cases, we derive a concentrated attack which maximizes the damage. □

*Corollary 1.* For any deterministic attack, and for any allocation and stochastic demand, concentrating the attack is an optimal attack strategy.

## 3.2 Concurrent Stochastic Attackers

While in the deterministic world each attacked resource fails, we now assume that the success of an attacking agent is stochastic: each attacked resource fails (independently) with probability $p$. Hence, the number of failures is a random variable with *Binomial* distribution and the number of surviving resources is $S_p(L_i, x_i) \sim L_i - \text{Bin}(x_i, p)$. This model is more general than the prior model and models realistic situations where the success of an attacking agent is a random variable, independent of the success of the other agents.

Now, the supply turns stochastic as well and the expression $\mathbb{E}[\min\{S_p(L, x), D\}]$ becomes more challenging as it forms a convolution. Yet, we prove that the marginal damage of the attacker remains monotonic under this settings.

*Theorem 2.* For any $D_i, L_i, i$ and $1 \leq x \leq L_i$,

$$\Delta_i(x) \leq \Delta_i(x+1).$$

We start by proving two lemmas and then conclude with proving the theorem.

*Lemma 1.* The marginal damage of attack agent $x$, $1 \leq x \leq L_i$, in region $i$ is:

$$\Delta_i(x) = r_i \cdot p \cdot \Pr[S(L_i, x-1) \leq D_i]. \quad (5)$$

PROOF. Let $G_i(x) = \min\{S_p(L_i, x), D_i\}$. By Eq. (1), we know that:

$$\Delta_i(x) = r_i \cdot (\mathbb{E}[G_i(x-1)] - \mathbb{E}[G_i(x)]).$$

Since the success of each agent is independent of the others, adding an attack agent to a region is beneficial if and only if: (1) The agent succeeds in failing a resource, (2) The corresponding resource is assigned to demand and failing it results in rejecting a request. Let $I_x$ be an indicator which gets the value of 1 if the $x$th attack agent successfully attacked a resource. Then,

$$G_i(x-1) - G_i(x) = \begin{cases} r_i & \text{if } S_p(L_i, x-1) \leq D_i \text{ and } I_x = 1 \\ 0 & \text{otherwise} \end{cases}.$$

Hence, $\Delta_i(x) = r_i \cdot p \cdot \Pr[S(L_i, x-1) \leq D_i]$, as required. □

*Lemma 2.* For any $d, L, x > 0$,

$$\Pr[S_p(L,x) \le d] \le \Pr[S_p(L,x+1) \le d].$$

PROOF. The proof follows from the inequality: $\Pr[\text{Bin}(x+1,p) \ge d] = \Pr[\text{Bin}(x,p) \ge d] + p \cdot (\Pr[\text{Bin}(x,p) \ge d-1] - \Pr[\text{Bin}(x,p) \ge d]) \ge \Pr[\text{Bin}(x,p) \ge d]$. For lack of space the full detailed proof is omitted. □

PROOF (THEOREM 2). We use Eq. (5) from Lem. 1 in order to prove the monotonicity of the marginal damage. Therefore, we need to prove that: $\Pr[S(L_i, x-1) \le D_i] \le \Pr[S(L_i, x) \le D_i]$. Using the *law of total probability* and Lem. 2 the inequality holds. □

Note that Thm. 1 in the previous section, stating that concentrating the attack is an optimal strategy under the deterministic model, was based on the monotonicity of the marginal damage function. Therefore, based on Thm. 2, Thm. 1 holds under this settings as well.

*Corollary 2.* For any binomial attack, and for any allocation and stochastic demand, concentrating the attack is an optimal attack strategy.

# 4. USER-MIGRATION DEFENSE

Some geographically distributed systems (such as cloud based server systems) can reduce the damage of attacks using dynamic user (request) migration. Under such mechanism the system may shift requests from one region to another based on dynamic fluctuations of the demand or of the resources availability. Thus, if one region experiences many failures due to an attack, the system can shift the requests of that region to another region, possibly incurring an additional cost. In this section we extend the analysis and study optimal attacks under request-migration architectures.

Consider a request made in the system. The request can be assigned to a resource located in the region where it is formed (satisfied *locally*). Otherwise, it can be assigned to a resource located in a remote region (satisfied *remotely*). A request is *unsatisfied* when it is not assigned to any resource in the system. There is a higher reward for satisfying a request locally (rather than remotely). Let $r$ be the reward for satisfying a request, and $r_i$ the bonus when satisfied locally.

The *assignment* problem (assigning a request to resources in the network, given a realization of the supply and demand) has a closed-form solution, as was proved in [2]. The idea is to start by assigning resources locally and then matching requests with remaining available remote resources. Hence, the system's revenue under an attack $\mathbf{X}$, $R(\mathbf{X})$, when using user-migration defence is:

$$\sum_{i=1}^{n} r_i \cdot \mathbb{E}[\min\{S(L_i,x_i), D_i\}] + r \cdot \mathbb{E}[\min\{\sum_{i=1}^{n} S(L_i,x_i), \|\mathbf{D}\|\}]. \quad (6)$$

In the following subsection, we study the effect of request-migration on the optimal attack strategy derived in Sec. 3.

## 4.1 Attacks under user migration

*Deterministic Attacks.* Let $\mathbf{L}$ and $\mathbf{D}$ be the system's allocation and demand, respectively. Let $\mathbf{X}_1$, $\mathbf{X}_2$ be attack vectors such that $\|\mathbf{X}_1\| = \|\mathbf{X}_2\|$. The supply $S$ of resources under a deterministic attack is: $S(L,x) = L - x$. Using Eq. (6), the system revenue under attack $\mathbf{X}_1$ is:

$$\sum_{i=1}^{n} r_i \cdot \mathbb{E}[\min\{L_i - x_i, D_i\}] + r \cdot \mathbb{E}[\min\{\sum_{i=1}^{n}(L_i - x_i), \|\mathbf{D}\|\}] =$$

$$= \sum_{i=1}^{n} r_i \cdot \mathbb{E}[\min\{L_i - x_i, D_i\}] + r \cdot \mathbb{E}[\min\{\|\mathbf{L}\| - \|\mathbf{X}_1\|, \|\mathbf{D}\|\}].$$

In Subsection 3.1 we proved that the first term of this equation optimizes when we concentrate the attacks. The second term of this equation is a function of $\|\mathbf{L}\|$, $\|\mathbf{D}\|$ and $\|\mathbf{X}_1\|$. According to our assumption, $\|\mathbf{X}_1\| = \|\mathbf{X}_2\|$. Therefore the second term has the same value under both attacks, since given $\mathbf{L}$ and $\mathbf{D}$ it depends only on the attack-size.

*Corollary 3.* For any deterministic attack, and for any allocation and stochastic demand, concentrating remains an optimal attack strategy under the user-migration defense.

*Concurrent Stochastic Attackers.* The supply $S_p$ under a binomial attack is $S_p(L,x) = L - Bin(x,p)$. The following is a known property of binomial distributions: For $\mathbf{X} = (x_1, \ldots, x_n)$ where $\forall i, x_i \ge 0$, $\sum_{i=1}^{n} \text{Bin}(x_i, p) = \text{Bin}(\|\mathbf{X}\|, p)$. Hence,

$$\min\{\sum_{i=1}^{n} S_p(L_i,x_i), \|\mathbf{D}\|\} = \min\{\|\mathbf{L}\| - \text{Bin}(\|\mathbf{X}\|,p), \|\mathbf{D}\|\}. \quad (7)$$

As in the deterministic model, the second term of this equation depends only on the attack-size, $\|\mathbf{X}\|$, and $p$, given $\mathbf{L}, \mathbf{D}$. In Subsection 3.2 we proved that the first term optimizes (given the attack size and success probability) when we concentrate the attack.

*Corollary 4.* For any binomial attack, and for any allocation and stochastic demand, concentrating remains an optimal attack strategy under the user-migration defense.

# 5. CONCLUSIONS AND FURTHER WORK

We analyzed optimal attacks on distributed resource systems with stochastic demand and showed that optimal attacks must be concentrated. This applies both to deterministic and binomial attacks. Furthermore, it holds also when systems use user-migration dynamic defense. The concentration property assists us in devising algorithms (which were not presented for lack of space) to derive the optimal attack on a given system. Further results of our work and an ongoing work develop the methodology to a wider set of attacks and study how to optimize resource placement against worst-case attacks.

# 6. REFERENCES


[1] P. Krishnan, D. Raz, and Y. Shavitt. The cache location problem. *IEEE/ACM transactions on networking, 2000.*, 8(5):568–582, 2000.

[2] Y. Rochman, H. Levy, and E. Brosh. Efficient resource placement in cloud computing and network applications. *SIGMETRICS Perform. Evaluation Rev., 2014.*, 42(2):49–51, 2014.

[3] E. Sherzer, G. Gilboa-Freedman, and H. Levy. Resource allocation in a cloud under virus attacks. In *Proceedings of the 11th EAI International Conference on Performance Evaluation Methodologies and Tools, VALUETOOLS 2017, Venice, Italy, 2017.*