

文章编号: 2096-1618(2017)02-0114-08

稿件来源: 科学网

网 址: <http://blog.sciencenet.cn/blog-453322-1005963.html>

发表时间: 2016-09-30

安全通论(16)

——黑客生态学

杨义先, 钮心忻

(北京邮电大学信息安全中心, 北京 100876)

摘要:影响网络空间安全的主角是黑客,控制住了黑客就掌握了安全。而控制黑客的最有效手段,就是控制黑客的生态环境;为此,需要首先设法了解这个生态环境。本文试图在最简单的情况下(即单种黑客工具),揭示黑客群体的诞生、发展、合作、竞争、迁移、死亡等生态环节的动力学特性,比如黑客数目或密度的解析公式、平衡态的局部或全局稳定性、周期系统的周期解的存在性和稳定性、持续生存性等。同时,还给出了控制黑客生态环境的一些建议,比如何时动手,何时放手,第一道防线设在哪,第二道防线设在哪等。关于较复杂的黑客、正常用户与红客相互作用的情形,将在后面的《安全通论》系列文章中单独研究。

doi: 10.16836/j.cnki.jcuit.2017.02.002

0 引言

关于黑客,外行看热闹,看到的是一个个绝顶聪明、行为怪诞的稀有动物;内行看门道,看到的却是一个个冷冰冰的黑客工具,因为,离开了工具,黑客就什么也不是了。所以,下面只关心黑客工具,当然,我们把工具也当作生物来描述,在不引起混淆的情况下,也用黑客之名来称呼。

除极个别顶级黑客会自己开发工具之外,绝大部分黑客,都只会使用现成的黑客工具(其实就是一些特殊软件)。而且,顶级黑客的杀手锏工具是决不外传的,所以,它不在本文的研究范围之内,让法律和红客去单挑这种顶级黑客吧。本文不研究的黑客工具还包括:预装类和广告类。比如,某些免费手机中已经悄悄预装了偷钱软件,这便是预装类的例子;某些靠广告支撑的畅销软件中的漏洞(有意或无意),便是广告类的例子。更严谨地说,本文只研究那些依靠口口相传,在网上广泛流行,并被普通黑客经常使用的黑客工具。这是因为,一方面,它们才是破坏力量的主体,虽然其媒体出境率并不高;另一方面,这类黑客工具的传播具有明显的生物特性,从而,可以借用现成的生物动力学成果。为简捷计,除非特别说明,本文后面所说的黑客工具,都限指这种依靠口口相传的黑客软件。

在现实中,一个黑客所拥有并随时使用的,肯定不止一种黑客工具,但是,为了研究方便,本文假定:所有黑客都只用同一种黑客工具。当然,这里所说的一种工具,也并非仅仅是一个工具,而至少是一类工具,比

如,若以黑客目标为准来分类的话,那么,所有试图获得正常用户的密码口令的工具,都可以当作一种工具。另外,我们说只有顶级黑客才会开发自己的工具,并不意味着普通黑客不对其工具做任何个性化的处理,但是,这种大同小异的修改,我们忽略不计。

1 单工具黑客动力学

当某种黑客工具,即一种软件,被开发出来并被放在网上后,还不能算作黑客就诞生了(最多只能算作黑客的首枚卵产出了),因为,没人用的软件等于不存在。只有当某人,下载并使用该软件去攻击别人时,才说一个黑客诞生了。这个黑客也许又会将该款软件推荐给他的一批朋友(相当于他又下了一批蛋),这批朋友中的某些人又去下载该软件并攻击别人,那么,又诞生了若干新一代,儿子代,黑客。这些新黑客又再向他们的朋友推荐,如此循环往复,于是,更新一代,孙子代,的黑客就源源不断地诞生了。你看,黑客的诞生模式,其实与鸡鸭猪狗等的诞生模式并无二异,都可用一棵树图来表示,该树图中的点就代表相应的黑客(或生物)。而且,最终的黑客总数会非常庞大,黑客代际之间的重叠会非常严重,以至于 t 时刻的黑客数目或密度 $N(t)$ 可以用 t 的连续函数来近似。

再确认一次,本文只考虑有同一种黑客软件的情形,相当于单种群的生物动力学。一来是因为,单工具黑客的研究相对容易,可以得到一些比较深入的结果;

二来因为,单工具黑客是多工具情形(相当于多种群生物)的基础;三来,如果被攻击的目标互不相关(比如,有的黑客是想获得隐私信息,有的黑客是想篡改别人的网页等),那么,就可以将这些黑客看作并列的几批使用单工具的黑客,从而,本文的所有成果对每批黑客都有效。

1.1 Malthus 模型

记 $N(t)$ 表示 t 时刻的黑客数目或密度,即正在用该工具攻击别人的黑客数量或密度(由于密度等于黑客数与总用户数的比值,所以在总数不变的情况下,密度和黑客数是等价的,不必刻意区分)。如果黑客的增长率是常数,或单位时间内黑客增长量与当时的黑客数量成正比,那么,就可用 b 和 d 分别表示黑客的出生率和死亡率(这里的所谓“死亡”,包括两大部分,其一,某人卸载了此软件,从而黑客总数减少一个;其二,某人虽拥有该工具,但是,此时此刻并未使用它去攻击别人,相当于生物的迁出,效果上也等于黑客数减少了一个。所谓“出生”,也包括两大部分,其一,就是某个新人下载并正使用该软件攻击别人,从而黑客总数增加一个;其二,前一时刻未出手的黑客,此刻发力了,相当于生物的迁入,效果上也等于黑客数增加了一个。)于是,在任意小的时间区间 Δt 内, $N(t)$ 的变化量满足等式: $N(t+\Delta t) - N(t) = bN(t)\Delta t - dN(t)\Delta t$ 。对该公式两边同除 Δt 并令 $\Delta t \rightarrow 0$ 取极限,得到了著名的 Malthus 微分方程

$\frac{dN(t)}{dt} = rN(t)$, 其中 $r = b - d$ 称为内禀增长率。

该微分方程的解析解为 $N(t) = N(0)e^{rt}$, 于是,根据 b 和 d 的大小,在 Malthus 模型下,黑客的最终数量将为 $\lim_{t \rightarrow \infty} N(t) = 0$, 当 $r < 0$ (即死亡率大于出生率时); $\lim_{t \rightarrow \infty} N(t) = N(0)$, 当 $r = 0$ (即死亡率等于出生率); $\lim_{t \rightarrow \infty} N(t) = \infty$, 当 $r > 0$ (即出生率大于死亡率)。由此可见,无论 r 多么小,只要 $r > 0$ (即出生率大于死亡率),那么,活跃黑客的最终数量将为无穷大,但是,实际情况显然不是这样的,因为,当黑客数量或密度大到某个程度后,合法用户的安全防护措施一定会加强,从而,使得该黑客工具失灵,导致黑客们不得不放弃该工具(转而寻求其他攻击手段),这相当于该黑客死亡,于是,死亡率会迅速超过出生率,黑客总数又会大减。

更准确地说, Malthus 模型仅仅适用于黑客工具刚刚出现的早期阶段,那时,黑客数量相对较少(或密度相对较小),红客的防护措施还比较薄弱,黑客攻击的成功率和利润都较高,从而,又会刺激更多的黑客诞生或迁入,即出生率增加,死亡率减少。但是,随着黑客

数量和密度的增大,觉醒并采取防卫措施的合法用户会增多,黑客的可攻击对象会减少,黑客彼此之间的竞争会加剧…,总之,死亡率增加,出生率减少,即内禀增长率减少,由此可见,不能永远假设 r 为常数,于是,便引出了下一小节(1.2节)的 Logistic 增长模型。

在单工具情形下,还需要引入另一个重要参数,称为黑客的最小生存数量(或密度),记为 K_0 ,它意指,如果黑客数 $N(t)$ 永远小于 K_0 时,那么,黑客数将逐步减少,并最终灭亡,即趋于 0。参数 K_0 的存在性可以这样来推理:由于黑客软件是(经朋友介绍后)自愿获取的,如果利用此工具去发动攻击会得不偿失,那么,他就会放弃该工具(即死掉一个黑客)并且不再向其朋友推荐;当越来越多的黑客死亡时,该种黑客工具便被淘汰了。相反,如果事实证明,该工具有利可图,那么,黑客就会继续拥有并使用该工具,并有可能向其朋友推荐,从而,黑客数将超过 K_0 。

在“不亏本”的前提下,人类本来就有相互合作的天性,特别是当 $N(t)$ 较小时,更会互相帮助(这便是“老乡见老乡,两眼泪汪汪”的人性依据),因为,帮助的结果对自己并无害(至少是害处很小),最终便导致提升黑客数量的增长率,甚至达到标准 Malthus 模型的指数增长速度。当然,当 $N(t)$ 较大时,情况就相反了,便会互相竞争(这便是“文人相轻”的人性依据),最终结果便是抑制黑客数量的增长率,这便是下一小节(1.2节)logistic 模型将要考虑的问题。

设 A 为黑客数的最大平均改变率,当 $N(t)$ 较小且 $N(t) > K_0$,生物学的经验已经告诉我们:黑客的内禀增长率 r 可以直观地替代为 $r = A[1 - K_0/N(t)]$, 于是,标准 Malthus 模型就变形为如下微分方程:

$$\frac{dN(t)}{dt} = AN(t)[1 - K_0/N(t)]$$

当 A 为正时,该微分方程存在零平衡态和正平衡态 K_0 ,而且,零平衡态是局部稳定的,即,当 $N(t) < K_0$ 后,黑客数 $N(t)$ 会不断减少,并最终趋于 0,于是该黑客工具被淘汰;但是,正平衡态 K_0 却是不稳定的,即,当 $N(t) > K_0$ 时,黑客数呈增加态势;而当 $N(t) < K_0$ 时,黑客数将不断减少并最终趋于 0。

上述分析对守护安全的红客们,有如下启发:(1)消灭黑客要宜早不宜迟,即,在黑客数还没有达到最小生存量 K_0 时就动手,效果最好;(2)如果成本较大,那么,不必对黑客斩尽杀绝,只需要将其问题控制在 K_0 之内,黑客便会自动灭亡;(3)如果错过了最佳时机(即黑客数已经超过 K_0),那么,黑客数将在随后的短时间内,呈现指数级的爆炸性增长,此时,不必与黑客硬拼,而应该充分运用黑客之间的竞争机制,让他们互相制约(见下面的 logistic 模型);(4)控制黑客的关键

是控制内禀增长率 r , 这又有两种思路: 其一是减少出生率 b ; 其二是增加死亡率 d 。如果能够使 $r < 0$, 那么, 就胜券在握; 如果能够使 $r = 0$, 那么, 就要考虑“任由 $N(0)$ 个黑客为非作歹”和“将黑客数量控制在 K_0 之内”的成本谁高谁低, 取成本低者而行动; 如果没法控制内禀增长率而出现了 $r > 0$, 那么, 红客的这第一道防线就崩溃了, 只能转战由 logistic 模型构建的下一道防线。

1.2 Logistic 增长模型及变形

每一款黑客工具都不可能永远通吃所有合法用户, 换句话说, 每个网络能够承受的活跃黑客数都是有限的, 该数称为环境容纳量, 记为 K (正数), 即, 当 $N(t) = K$ 时, 黑客数将出现零增长 (当然, 不难看出, 一定有 $K_0 < K$)。其实, 在实践中, 往往不是黑客工具有多么厉害, 而是合法用户太懒或太大意, 比如, 他们懒于安装相关的漏洞补丁或缺乏安全意识等; 但是, 一旦活跃黑客数量或密度过大, 以致在某合法用户身边出现了受害者时, 他就会积极加强防护, 从而, 扼制了黑客有生存环境。

黑客数的内禀增长率当然不会突然陡减为 0, 合理的假定是: 随着活跃黑客数逐渐靠近环境容量 K 时, r 逐渐变小并最终靠近 0。最简单的情况是: 每增加一个黑客, 就均匀地对内禀增长率产生 $1/K$ 抑制影响; 于是, $N(t)$ 个黑客就产生 $N(t)/K$ 的抑制影响, 从而, 未被影响的部分就剩下 $1 - N(t)/K$, 换句话说, 内禀增长率就由 r 减少为 $r[1 - N(t)/K]$, 于是, 内禀增长率为常数的 Malthus 模型, 便被改进为内禀增长率为变数 $r[1 - N(t)/K]$ 的如下微分方程所表示的标准 Logistic 模型:

$$dN(t)/dt = rN(t)[1 - N(t)/K],$$
 其中 r 是内禀增长率, K 为环境容纳量

该微分方程的解析解为 $N(t) = KN(0) / \{N(0) - [N(0) - K]e^{-rt}\}$, 它完全由 r 、 K 、和黑客数量的初值 $N(0)$ 确定。根据此解析解, 我们得知:

若 $N(0) > 0$ 时, 当 $t \rightarrow \infty$, 黑客数 $N(t)$ 将最终趋于容纳量 K ;

而且, 当初值 $N(0)$ 满足 $0 < N(0) < K/2$ (即, 黑客初值数不超过容纳量的一半) 时, 黑客数量的曲线 $N(t)$ 将呈现 S 型; 并且在 $K/2$ 点处, 出现唯一的拐点: 当 $N(t)$ 很小时, 在一定的时间范围内, 黑客数将成指数增长模式; 然后, 抑制影响开始发挥作用, 并在容纳量 K 处, 黑客数量将最终达到饱和。更详细地说, 此处的 S 曲线, 可以划分为五个阶段: (1) 开始期, 也称为潜伏期, 黑客数量很少, 数量和密度的增长缓慢; (2) 加速

期, 随着黑客数的增加, 密度也迅速增加; (3) 转折期, 当黑客数达到饱和密度的一半 ($K/2$) 时, 密度增长最快; (4) 减速期, 当黑客数超过 $K/2$ 以后, 密度增长逐渐变慢; (5) 饱和期, 黑客数量达到 K 值而饱和, 这意味着 K 是稳定的。

上述标准的 logistic 模型更适合于黑客数量和密度 $N(t)$ 较大时的情况, 它已经考虑到了黑客彼此之间的竞争, 以及由此导致的对内禀增长率的抑制情况。而当 $N(t)$ 较小时, 黑客之间又是相互帮助的, 并将导致内禀增长率变大, 所以, 若同时考虑“人少时的合作”和“人多时的竞争”, 那么, 标准 logistic 模型便可改进为如下“具有 Allee 效应的 logistic 模型”:

$$dN(t)/dt = rN(t)[N(t)/K_0 - 1][1 - N(t)/K]$$

此时, 便存在着三个非负平衡态: 0 、 K_0 和 K 。具体地说:

当 $0 < N(t) < K_0$ 时, $dN(t)/dt < 0$, 即, 黑客数量不断减少;

当 $K_0 < N(t) < K$ 时, $dN(t)/dt > 0$, 即, 黑客数量不断增加;

当 $N(t) > K$ 时, $dN(t)/dt < 0$, 即, 黑客数量又不断减少。

因此, 0 和 K (最大容纳量) 是局部稳定的平衡态; 黑客的最小生存数量 K_0 是不稳定的平衡态, 并且它有两个稳定平衡态的分界点, 即, 当黑客数量的初值 $N(0) > K_0$ 时, 黑客数量将最终趋于 K ; 而当 $N(0) < K_0$ 时, 黑客数将最终趋于零, 该黑客工具被淘汰。

除了考虑黑客合作时的改进型 logistic 模型 (即具有 Allee 效应的 logistic 模型) 之外, 还可以再考虑正常用户合作时的改进 logistic 模型。此时, 当某个用户被攻击后, 他不但会自身加强保护措施, 还会将其经验和教训传播给身边人员, 提醒他们注意, 于是, 黑客可能攻击的对象数就会减少, 形象地说, 黑客的“食物”就减少了。极端情形是: 如果所有用户都觉悟并采取防护措施后, 那么, 该黑客工具就失灵了, 从而, 黑客就无目标可攻击, 当然也就只好灭亡了。

记 S 为黑客数达到饱和时, 正常用户的觉悟率 (即, 他们采取了安全措施, 使得该款黑客工具失效); 记 $F(t)$ 为 t 时刻 (黑客数为 $N(t)$ 时) 的用户觉悟率。将标准 logistic 方程等价地重新写为 $(1/N(t)) [dN(t)/dt] = r[K - N(t)]/K$, 若保持该公式的左边不变, 但将其右边的“饱和量 K ”替换为“饱和时的用户觉悟率 S ”, 将右边的“黑客数 $N(t)$ ”替换为“用户觉悟率”, 于是, 便得到标准 logistic 模型的如下另一种改进“用户合作时的 logistic 模型”:

$$(1/N(t)) [dN(t)/dt] = r[S - F(t)]/S$$

它的左边表示“ t 时刻,黑客的平均增长率”;而它的右边则表示“ t 时刻,用户的未觉悟率”。该公式的直观解释便是:“黑客增长率”与“用户未觉悟率”成正比。这种解释显然是有道理的,因为,未觉悟的用户越多,黑客的利润就越大,就越能刺激更多的黑客发动进攻;反之亦然。

再注意到 $F(t)$ 当然应该与黑客 $N(t)$ 和黑客的变化数 $dN(t)/dt$ 有关,为简便计,假定这种关系是线性关系,即, $F(t) = c_1 N(t) + c_2 dN(t)/dt$, 这里 $c_1, c_2 > 0$, 即,黑客越多,黑客增长越快,那么,觉悟的用户也会更快地增长。由于在饱和状态时,同时成立 $dN(t)/dt = 0$ 、 $N(t) = K$ 和 $F(t) = S$, 所以,在公式 $F(t) = c_1 N(t) + c_2 dN(t)/dt$ 中,让时间趋于无穷大后,便有 $S = c_1 K$ 。于是,上面的“用户合作时的 logistic 模型”便可以更具体地表述为:

$$dN(t)/dt = rN(t) [K - N(t)] / [K + rcN(t)],$$
 这里 $c = c_2/c_1$

该微分方程的解析解为 $N(t) = Ae^{rt} [|K - N(t)|^{1+rc}]$, 其中 A 是由初始条件确定的常数。注意到当时间趋于无穷时,左边为有限;而右边的 e^{rt} 为无穷大,所以要使整个右边有限的话,就必须有 $|K - N(t)|^{1+rc}$ 趋于 0, 即 $N(t)$ 趋于 K 。

该微分方程还可看出:当黑客数 $N(t)$ 较小时,黑客数的增加,反而会使得黑客的增长率 $dN(t)/dt$ 减少;当黑客较大时,黑客数的增加才会同时促进黑客增长率也增加。这再一次印证了:消灭黑客宜早不宜迟。

此小节的上述分析,可以给红客以如下启发:(1) 如果治理黑客的成本高于“任由 K (容纳量) 个黑客肆虐”的成本,那么,就不必治理了,否则就是吃力不讨好;(2) 如果未能在开始期消灭黑客(即设置在 K_0 处的第一道防线被突破),那么,第二道最佳防线就应该设置在 $N(t) = K/2$ 处的转折期;(3) 如果第二道防线也被突破了,那么,就应该重点保护关键用户,不必再设置第三道防线了,除非有特殊的非经济因素;(4) “用户彼此合作,提升觉悟率”也是对付黑客的另一个有效手段。

1.3 非自治单工具模型

标准 Logistic 模型的一个重要假设就是:内禀增长率 r 和容纳量 K 均为常数。这种假设的优点是:直观简洁且逼近实际。当然,严格说来, r 和 K 不会永远都是常数,也会变化,比如,当黑客的期望值变大时,更多的黑客将因无利可图而放弃攻击(当然也就放弃工具了),那么,容纳量将变小;当合法用户变得更麻木时,黑客能够获得的利润将更多,从而,将有能力滋养更多

的黑客,即,容纳量就会增大。不过,每一种模型都有不够精确的地方,我们必须在取舍之间寻找折中,毕竟当模型过于精细后,相应的微分方程就无法求解了,更不能为了精细而精细。

若将 r 和 K 分别用分段连续的时间函数 $r(t)$ 和 $K(t)$ 来替代,那么,标准 logistic 模型就变成了如下非自治的 logistic 模型:

$$dN(t)/dt = r(t)N(t) [1 - N(t)/K(t)]$$

该微分方程的解析解为:

$$N(t) = N(0) \exp \left[\int_0^t r(s) ds \right] / \left\{ 1 + N(0) \int_0^t \exp \left(\int_0^s r(f) df \right) r(s) / K(s) ds \right\}$$

如果 $0 < \inf_{t>0} r(t) \leq r(t) \leq \sup_{t>0} r(t) < \infty$ 并且 $0 < \inf_{t>0} K(t) \leq K(t) \leq \sup_{t>0} K(t) < \infty$, 那么,非自治的 logistic 模型就有一个全局稳定的解:

$$N^*(t) = \left\{ \int_0^\infty \exp \left[- \int_0^s r(t-f) df \right] r(t-s) / K(t-s) ds \right\}^{-1}$$

并且,当 $r(t)$ 和 $K(t)$ 是周期函数时, $N^*(t)$ 也是周期的。下面再进一步地,分成一些特殊情况,来讨论非自治 logistic 模型:

情况 1 环境退化

所谓环境退化,就是指黑客的生存条件越来越差,即,黑客的容纳量 $K(t)$ 虽非负,但随着时间的推进 $K(t)$ 越来越小,甚至 $\lim_{t \rightarrow \infty} K(t) = 0$ 。此时已经证明:

如果内禀增长率满足 $\int_0^\infty r(s) ds = \infty$, 则 $\lim_{t \rightarrow \infty} N(t) = 0$ 。

该结果的直观解释便是:即使内禀增长率较大,在退化环境下,随着时间的推移,该黑客工具也将最终被淘汰,黑客被消灭。

如果内禀增长率满足 $\int_0^\infty r(s) ds < \infty$, 则 $\lim_{t \rightarrow \infty} N(t) = N_\infty < \infty$ 是一个正常数。由此,对比上面那个直观解释,我们就得到一个有趣结果:即使内禀增长率较小,黑客数也会长期维持在正常数 N_∞ 附近,它与初始值无关。由于内禀增长率是在无外界影响的条件下,黑客数量的自然增长率(这便是“内禀”的含义所在),由此(再结合标准 logistic 模型的结果)可知:如果某种黑客工具的内禀增长率很低,那么,它在非常有利的环境下可能也很难生存;但是,在退化的环境下,它却可能长期生存甚至繁荣!

情况 2 周期性的考虑

黑客世界中也有一些有趣的周期现象,比如,在无外界干扰时,从宏观上看,当内禀增长率 $r(t)$ 越来越大时,黑客的数量会增多,因此,每个黑客的利润会越来越少,这就会反过来促进越来越多的黑客放弃攻击,从而使 $r(t)$ 开始变小;换句话说, $r(t)$ 会不断地周期性振荡。同样,容纳量 $K(t)$ 也具有这种周期特性。为数学上处理方便,我们干脆假设 $r(t)$ 和 $K(t)$ 就是周期为

T 的连续函数,并且,还做出如下三个合理的假设:

假设 1 黑客数越来越多时,他们会彼此竞争,从而会越来越严重地抑制黑客数量的增长;

假设 2 当黑客数超过一定的值后,平均到每个黑客的利润会越来越低,因此,黑客数目将不会增加;

假设 3 在一个周期里,内禀增长率是受控的,即, $0 < \int_0^T r(t) dt < \infty$ 。

于是,此时非自治的 logistic 模型微分方程 $dN(t)/dt = r(t)N(t)[1-N(t)/K(t)]$ 存在着周期解析解: $N(t+T) = N(t)$, $N(0) = N(T)$ 为黑客的初始值,并且当 $0 < t < T$ 时,

$$N(t) = \left[\exp \left(\int_0^T r(f) df - 1 \right) \left\{ \int_t^{t+T} [r(s) \exp \left[- \int_s^{t+T} r(f) df \right] / K(s)] ds \right\}^{-1} \right]$$

情况 3 时滞因素的考虑

在标准 logistic 模型中, t 时刻黑客数的平均变化率 $(1/N(t)) [dN(t)/dt]$ 只与该时刻的黑客数有关,即,等于 $r[1-N(t)/K]$ 。但是,如果考虑得更精细一点,将会发现,其实存在着某种时滞现象,即, t 时刻黑客数的平均变化率,应该与 $t-\tau$ 时刻的黑客数有关,于是,便有标准 logistic 模型便可以改进为 $(1/N(t)) [dN(t)/dt] = r[1-N(t-\tau)/K]$, 或者,等价地就有带时滞的 logistic 模型:

$$dN(t)/dt = rN(t)[1-N(t-\tau)/K]$$

它存在着零平衡态,并且,当 $r > 0$ 时,零平衡态是不稳定的。此外,它还有一个正平衡态 $N=K$,其稳定性为:当 $0 \leq r\tau < \pi/2$ 时,平衡态 $N=K$ 是渐近稳定的;当 $r\tau > \pi/2$ 时,平衡态 $N=K$ 不稳定,此时,黑客 $N(t)$ 存在一个周期解,即,黑客数的变化呈周期性起伏。

此小节的上述分析,可以给红客以如下启发:(1) 如果能够控制黑客的生存环境,那么,增长态势越猛的黑客工具可能越短命;而增长缓慢的黑客工具可能会更命长,不过,如果他们的危害不高于治理成本的话,那么,就可以不理他们;(2) 黑客的增长率、网络对黑客数量的容量值、黑客数等都可能呈现出周期起伏的现象,因此,如果红客要想稳准狠地消灭黑客的话,最好在这些周期的低潮时下手!

2 单工具随机模型

上一节在研究黑客动力学时,忽略了所有随机因素。但是,在实际情况下,随机因素显然是存在的,因此,本节就来重点考虑随机性。

为减轻阅读负担,上一节我们几乎省略了所有复杂的数学推导。这是因为,虽然微分方程的求解很难,

但是,给出解析解后,验证其正确性却很容易,所以,我们的省略一点也没影响文章的严谨性和正确性,只是把大量的推导工作隐没在了后台而已。但是,本节中有些数学推导就无法省略了,希望这些必不可少的公式,不会给读者增添过多的困难。

2.1 纯生过程

这里的所谓“纯生”,就是假定没有死亡(含迁出,下同),即黑客只增不减。记 t 时刻黑客数为 $N(t)$,并假定:

(1) 每个黑客的诞生(含迁入,下同)是互相独立的;

(2) 在任意小的时间段 Δt 内,每个黑客诞生一个新黑客的概率为 $\lambda \Delta t + o(\Delta t)$,没有新黑客诞生的概率为 $1 - \lambda \Delta t + o(\Delta t)$,多于一个新黑客诞生的概率为 $o(\Delta t)$ 。

如果已知 $N(t) = n$,那么,在区间 $(t, t+\Delta t]$ 内诞生的新黑客个数,服从参数 n 和 $\lambda \Delta t$ 的二项分布的随机变量。当 Δt 非常小时,可以忽略 $o(\Delta t)$ 的影响。于是,当 $k=0, 1, \dots, n$ 时,有

$$P\{k \text{ 个新黑客在区间 } (t, t+\Delta t] \text{ 诞生} \mid N(t) = n\} = C(n, k) (\lambda \Delta t)^k (1 - \lambda \Delta t)^{n-k}$$

记该概率为 $P(k)$,这里和今后 $C(n, k)$ 都表示组合数公式,即 $C(n, k) = n! / (k! (n-k)!)$ 。于是, $P(0) = (1 - \lambda \Delta t)^n = 1 - \lambda n \Delta t + o(\Delta t)$; $P(1) = \lambda n \Delta t (1 - \lambda \Delta t)^{n-1} = \lambda n \Delta t + o(\Delta t)$;并且,当 $k \geq 2$ 时, $P(k) = o(\Delta t)$ 。换句话说,这意味着随机过程 $N = \{N(t), t \geq 0\}$ 是一个连续时间 Markov 过程。记 $N(0) = a > 0$,现在考虑黑客数的转移概率 $p_n(t) = P(N(t) = n \mid N(0) = a)$, $a > 0, t > 0$,它显然只依赖于时间差,从而是一个平稳随机过程。

现在考虑 $p_n(t)$ 和 $p_n(t+\Delta t)$ 的关系。如果 $N(t+\Delta t) = n > a$ 且当 $\Delta t \rightarrow 0$ 时,忽略多于一个新黑客诞生的可能性,那么,在 t 时刻, $N(t)$ 就满足: $N(t) = n$,若在 $(t, t+\Delta t]$ 时间段内没有新黑客诞生; $N(t) = n-1$,若在 $(t, t+\Delta t]$ 时间段内有一个新黑客诞生。应用全概率公式,便有:

$$p_n(t+\Delta t) = (1 - \lambda n \Delta t) p_n(t) + \lambda (n-1) \Delta t p_{n-1}(t) + o(\Delta t), n > a$$

将该公式等价地变形为

$$[p_n(t+\Delta t) - p_n(t)] / \Delta t = \lambda [(n-1) p_{n-1}(t) - n p_n(t)] + o(\Delta t) / \Delta t$$

在该式中,令 $\Delta t \rightarrow 0$ 取极限,便有

$$dp_n(t)/dt = \lambda [(n-1) p_{n-1}(t) - n p_n(t)], n = a+1, a+2, \dots$$

当 $n=a$ 时,由于此前黑客数为 $a-1$ 的概率为 0 (因为纯生),所以,由全概率公式就有 $P\{N(t+\Delta t) =$

$a\} = P\{N(t+\Delta t) = a \mid N(t) = a\} P\{N(t) = a\}$, 所以

$$dp_a(t)/dt = -\lambda ap_a(t)$$

求解此微分方程后, 就有 $p_a(t) = e^{-\lambda at}$, 据此和前面已有的公式 $dp_n(t)/dt = \lambda[(n-1)p_{n-1}(t) - np_n(t)]$, 我们可以得到, 在 t 时刻, 有 k 个新黑客诞生的概率为

$$p_{a+k}(t) = C(a+k-1, a-1) e^{-\lambda at} (1-e^{-\lambda t})^k,$$

这里 $k=0, 1, 2, \dots, n$, 并且 $C(m, n)$ 为组合数公式。提醒: 这个公式实际上就给出了在“0 时刻黑客数为 a ”的条件下, t 时刻的黑客数达到 $a+k$ 的概率 $p_{a+k}(t)$; 因此, 在该时刻黑客数的均值 $\mu(t)$ 就为

$$\mu(t) = E(N(t) \mid N(0) = a) = \sum_{n=a}^{\infty} np_n(t) = ae^{\lambda t}$$

此处前两个等式来源于均值的定义和 $p_{a+k}(t)$ 的表达式, 最后一个公式中略去了详细的计算过程(见文献[17]的5.3.1节)。这个公式告诉我们一个有趣的结果: 在纯生过程中, t 时刻黑客的平均个数为 $ae^{\lambda t}$, 它与出生率为 $b=\lambda$ 的 Malthus 模型的解析式完全一样! 仔细想来也是有道理的, 因为, Malthus 模型更适用于黑客数(密度)较小的初期, 此时死亡(放弃工具)和迁出(有工具却不用)的黑客几乎不存在, 这当然可以看作一个纯生过程了。

2.2 纯灭过程

与纯生相反, 此时只有死亡(放弃或不用黑客工具)。假定某黑客 t 还存活但在时间段 $(t, t+\Delta t]$ 内死亡的概率为 $\mu\Delta t + o(\Delta t)$, 现在考虑条件转移概率

$$p_n(t) = P(N(t) = n \mid N(0) = a), n = a, a-1, \dots, 2, 1, 0$$

先看一个特殊情况 $a=1$, 那么, $p_1(t)$ 就是单个黑客在 t 时刻仍然存活的概率, 并且有

$$p_1(t+\Delta t) = p_1(t)(1-\mu\Delta t) + o(\Delta t)$$

其中 $1-\mu\Delta t$ 是单个黑客在时间段 $(t, t+\Delta t]$ 内没有死亡的概率。令 $\Delta t \rightarrow 0$ 取极限, 便有如下微分方程 $dp_1(t)/dt = -\mu p_1(t)$, $t > 0$, 它对初值 $p_1(0) = 1$ 的解为 $p_1(t) = e^{-\mu t}$ 。

如果在初始时刻的黑客数 $a > 1$, 则在 t 时刻仍然存活的黑客数是一个服从参数 a 和 $p_1(t)$ 的二项分布的随机变量, 所以有

$$p_n(t) = C(a, n) e^{-\mu t} (1-e^{-\mu t})^{a-n}, \quad n = a, a-1, \dots, 2, 1, 0$$

其相应的数学期望值和方差分别是

$$E[N(t)] = ae^{-\mu t} \text{ 和 } \text{Var}[N(t)] = ae^{-\mu t} [1-e^{-\mu t}]$$

可见, 此时黑客数量的变化规律与 Malthus 增长模型中 $d=\mu, b=0$ (有死无生) 的情形相似。

在纯灭过程中, 黑客数要么保持常数, 要么递减, 最终有可能变为 0 (即, 灭绝)。精确地说, 这种黑客工具灭绝的概率为

$$p_0(t) = P(N(t) = 0 \mid N(0) = a) = (1-e^{-\mu t})^a \rightarrow 1, t \rightarrow \infty$$

换句话说, 此时黑客灭绝的概率为 1, 一定灭亡。

2.3 线性出生和死亡的生灭过程

现在考虑同时有生也有死的情况, 为简单计, 假设生死速度均为线性。

设初始黑客数为 a 且在时刻 t 时, 黑客个数为 $N(t)$, 在时间区间 $(t, t+\Delta t]$ 内有一个新黑客诞生的概率为 $\lambda\Delta t + o(\Delta t)$, 有一个黑客死亡的概率为 $\mu\Delta t + o(\Delta t)$ 。于是, 在“ $N(t) = n$ ”的条件下, 在区间 $(t, t+\Delta t]$ 内出生一个黑客的概率为 $\lambda n\Delta t + o(\Delta t)$; 死亡一个黑客的概率为 $\mu n\Delta t + o(\Delta t)$; 黑客数不变的概率为 $1 - (\lambda + \mu)n\Delta t + o(\Delta t)$ 。所以, 仿前面, 记 $p_n(t) = P(N(t) = n \mid N(0) = a)$, 那么, 利用全概率公式, 便有

$$p_n(t+\Delta t) = p_{n-1}(t)\lambda(n-1)\Delta t + p_n(t)[1 - (\lambda + \mu)n\Delta t] + p_{n+1}(t)\mu(n+1)\Delta t + o(\Delta t)$$

该式两边同除 Δt , 并令 $\Delta t \rightarrow 0$, 于是, 在 $n \geq 1$ 时便得微分方程

$$dp_n(t)/dt = \lambda(n-1)p_{n-1}(t) - (\lambda + \mu)np_n(t) + \mu(n+1)p_{n+1}(t)$$

若 $n=0$, 则有 $dp_0(t)/dt = \mu p_1(t)$, 相应的初始条件为若 $n=a$, 则 $p_n(0) = 1$; 若 $n \neq a$, 则 $p_n(0) = 0$ 。

至此得到了有生有死情况下, 黑客个数的随机过程 $p_n(t)$ 所应该满足的微分方程, 由于求解此方程很复杂, 我们只给出最终结果如下:

记 $\Phi(s, t) = \sum_{n=0}^{\infty} p_n(t)s^n$, 于是 $p_n(t)$ 就是函数 $\Phi(s, t)$ 关于参量 s 的多项式展开式中 s^n 的系数。根据文献[17]的第5.3节的结果, 我们知道,

当 $\lambda \neq \mu$ 时, $\Phi(s, t) = \{[\mu - \psi(s)e^{-(\lambda-\mu)t}]/[\lambda - \psi(s)e^{-(\lambda-\mu)t}]\}^a$, 其中 $\psi(s) = (\lambda s - \mu)/(s-1)$;

当 $\lambda = \mu$ 时, $\Phi(s, t) = \{[1 - (\lambda t - 1)(s-1)]/[1 - \lambda t(s-1)]\}^a$ 。

现在来分析黑客被灭绝的概率, 即 $p_0(t) = P(N(t) = 0 \mid N(0) = a)$, 它其实就是 $\Phi(0, t)$, 所以,

当 $\lambda \neq \mu$ 时, $p_0(t) = \Phi(0, t) = \{[\mu(1 - e^{-(\lambda-\mu)t})]/[\lambda - \mu e^{-(\lambda-\mu)t}]\}^a$ 。若更进一步分析, 当 $\lambda < \mu$ 时, 在此式中令 $t \rightarrow \infty$, 那么就有 $p_0(t) \rightarrow 1$, 即该种黑客工具以概率 1 被灭绝(这是可以理解的, 因为, 新黑客出生的概率小于死亡概率时, 当然最终会灭绝); 当 $\lambda > \mu$ 时, 在此式中令 $t \rightarrow \infty$, 那么就有 $p_0(t) \rightarrow (\mu/\lambda)^a$, 即该种黑客数量将最终稳定在 $(\mu/\lambda)^a$ 。

当 $\lambda = \mu$ 时, $p_0(t) = \Phi(0, t) = [\lambda t / (\lambda t + 1)]^a$, 而且, 当 $t \rightarrow \infty$ 时, 也有 $p_0(t) \rightarrow 1$, 即该种黑客以概率 1 被灭绝。初看起来, 当出生概率等于死亡概率时, 好像很

难理解为什么它一定会灭绝。其实仔细分析后,就知道:0(灭绝)是一个吸引状态,且与 $N(t)$ 的距离是有限的,又由于黑客数量的轨迹的随机性,因此,掉进吸引子(灭绝)当然就成为必然了。

经认真计算后,还知道在此处有生有死的情况下,在黑客数初值为 $N(0)=a$ 的条件下, $N(t)$ 的条件数学期望值(即平均黑客数)

$$E[N(t) | N(0)=a] = ae^{(\lambda-\mu)t}$$

它与确定性的 Malthus 模型的增长情况一样。

在黑客数初值为 $N(0)=a$ 的条件下, $N(t)$ 的条件方差值为:

$$\text{当 } \lambda \neq \mu \text{ 时, } \text{Var}[N(t) | N(0)=a] = a(\lambda + \mu) e^{(\lambda-\mu)t} [e^{(\lambda-\mu)t} - 1] / (\lambda - \mu)$$

$$\text{当 } \lambda = \mu \text{ 时, } \text{Var}[N(t) | N(0)=a] = 2a\lambda t。$$

2.4 非自治线性生灭过程

小节(2.3)考虑出生概率和死亡概率时,故意忽略了时间和黑客数,其实黑客数越多时,其出生和死亡的概率也就越大,因此,更精细地假设:在 t 时刻,当黑客数为 n 时,相应的出生概率为 $\lambda_n = \lambda(t)n$ 和死亡概率为 $\mu_n = \mu(t)n$,于是,类似地可知条件概率 $p_n(t) = P(N(t)=n | N(0)=a)$ 满足如下微分方程:

$$dp_n(t)/dt = -n[\lambda(t) + \mu(t)]p_n(t) + (n-1)\lambda(t)p_{n-1}(t) + \mu(t)(n+1)p_{n+1}(t), n=1,2,\dots,n$$

和 $dp_0(t)/dt = \mu(t)p_1(t)$, $p_a(0)=1$, 且当 $n \neq a$ 时有 $p_n(0)=0$ 。

并最终求出(详见文献[17]的5.3.5节),在初始黑客数为 a 的条件下, t 时刻黑客数 $N(t)$ 的数学期望值为: $E[N(t) | N(0)=a] = a \exp \{ \int_0^t [\lambda(s) - \mu(s)] ds \}$ 。

2.5 增长率只与黑客有关的情况

假如黑客的出生率和死亡率都只与黑客个数有关,而与时间无关,不妨记:当有 n 个黑客时,出生率和死亡率分别为 λ_n 和 μ_n ; $N(t) \in \{0,1,2,\dots,K\}$ 为 t 时刻的黑客个数,那么,与前面类似,在 t 时刻, $N(t)$ 满足如下 Markov 方案

$$P\{N(t+\Delta t) = n+1 | N(t) = n\} = \lambda_n \Delta t + o(\Delta t),$$

$$P\{N(t+\Delta t) = n-1 | N(t) = n\} = \mu_n \Delta t + o(\Delta t),$$

$$P\{N(t+\Delta t) = n | N(t) = n\} = 1 - (\lambda_n + \mu_n) \Delta t + o(\Delta t),$$

令 $\Delta t \rightarrow 0$,取极限便得到黑客数的条件转移概率所满足的如下两个微分方程:

$$dp_0(t)/dt = \mu_1 p_1(t)$$

$$dp_n(t)/dt = \lambda_{n-1} p_{n-1}(t) - (\lambda_n + \mu_n) p_n(t) + \mu_{n+1} p_{n+1}(t)$$

求解这个微分方程很难,不过幸好我们需要的有关黑客何时会灭绝的结果可以描述如下(见文献[17]

的5.4节):

所谓灭绝时间,就是指当黑客数首次为0的时间,也可以理解为这种黑客在被最终淘汰前的持续时间。若记 T_n 为初始黑客数为 n 的情况下,黑客被灭绝的时间,它显然是一个随机变量,不过,该随机变量的均值为:

$$E[T_n] = \sum_{i=1}^n \sum_{j=i}^K (1/\mu_j) \prod_{h=i}^{j-1} (\lambda_h/\mu_h)$$

形象地说,对 $E[T_n]$ 越小的黑客工具,其寿命就越短。

3 结束语

在对付病虫害的长期过程中,人类已经知道:直接灭虫只是治标;控制害虫的生态环境,才是治本。对付网络黑客其实也是这个道理,但是,由于黑客们来无踪,去无影,所以,要想完全搞清其生态链,绝非易事。本文虽然在一些(比较合理的)人为假设下,揭示了黑客的部分生态特性,但是,大家别太乐观,因为,万里长征才迈出第一步,有待解决的问题还很多,比如:如何用实测数据来验证相关模型的逼真程度?(这需要大型,甚至国家级的,安全监测机构的数据支持;普通用户无能为力)各种模型中的相关参数如何来确定,现有的参数回归方法是否有效?模型是否能够进一步优化?如何利用已知的黑客生态学结果,去完成某款黑客工具的实际控制?等等。

另一方面,读者也别太悲观。因为,虽然黑客生态很复杂,但是,我们可以站在巨人肩上,借鉴生物数学家们过去一百多年来积累的众多成果。坦率地承认,目前生物学家的这些成果,对网络安全专家来说,还有点像天书。幸好安全专家都是善于攻坚克难的跨界精英,相信在不远的将来,一定会把天书破译。随着黑客生态学研究的逐步深入,必有更多的秘密被发现。假如有幸能把某些生物学家吸引到网络安全领域来,那么,跨学科的合作将如虎添翼。

笔者创立《安全通论》的最终目标是,为网络安全学科的各分支,建立统一的基础理论。因此,本系列的每篇文章,都必须站得安全界的最高点,以便从宏观上研究普遍的共性问题,而忽略细枝末节。但是,我们的不少方法和思路,其实都可用于解决一些更具体的问题。读者若意识到这一点,也许会有助于在自己的安全分支中,获得意外丰收。比如,本文的成果几乎可以完全照搬地应用于电信诈骗、非法传销、网络欺诈等具体安全问题的治理。而且,在解决电信诈骗等问题时,描述其安全生态的检测数据更容易获得,相关模

型更容易验证,相关参数也更容易确定。甚至,这些局部成果可能会反过来,帮助建立通用的黑客生态学。不过,为了保持《安全通论》结构的整洁性,我们就不纠缠这些细节了。欢迎有特殊兴趣的读者将《安全通论》应用于解决任何具体的安全问题,我们将毫无保留地,提供力所能及的帮助。

本文多次强调“单工具”,是想避免陷入不必要的细节纠纷。其实,假如有多种黑客工具,那么,由于每种工具的传播都具有生物繁殖特性,所以,粗略地说,各位黑客汇聚在一起时,也可以看作一类“单种群生物”,从而,本文的所有思路和结果都仍然有效。当然,如果把黑客、正常用户和红客放到一起时,就不能再把他们看成同一个“物种”了,毕竟他们彼此之间的对抗多于协作。

包括本文在内的《安全通论》系列论文,还特别注意尽量不越界,即,始终以安全为对象。其实,我们的许多思路和结果也可以在安全之外的领域发挥作用,比如,其实除了普通的黑客软件之外,包括微信、高德地图等在内的几乎所有 App 和其他非预装类软件,都具有本文揭示的共同生物繁殖特性,所以,关于它们的生态学特性,完全可以借用本文的结果。这显然对这些软件经营的商业模式、用户分布、升级维护和产品推广等方面都是很有帮助的。

参考文献:

- [1] 杨义先,钮心忻. 安全通论(1)之“经络篇”[EB/OL]. <http://blog.sciencenet.cn/blog-453322-944217.html>, 2015-12-18.
- [2] 杨义先,钮心忻. 安全通论(2):攻防篇之“盲对抗”[EB/OL]. <http://blog.sciencenet.cn/blog-453322-947304.html>, 2016-01-01.
- [3] 杨义先,钮心忻. 安全通论(3):攻防篇之“非盲对抗”之“石头剪刀布游戏”[EB/OL]. <http://blog.sciencenet.cn/blog-453322-948089.html>, 2016-01-04.
- [4] 杨义先,钮心忻. 安全通论(4):攻防篇之“非盲对抗”之“童趣游戏”[EB/OL]. <http://blog.sciencenet.cn/blog-453322-949155.html>, 2016-01-09.
- [5] 杨义先,钮心忻. 安全通论(5):攻防篇之“非盲

对抗”收官作及“劝酒令”[EB/OL]. <http://blog.sciencenet.cn/blog-453322-950146.html>, 2016-01-13.

- [6] 杨义先,钮心忻. 安全通论(6):攻防篇之“多人盲对抗”[EB/OL]. <http://blog.sciencenet.cn/blog-453322-954445.html>, 2016-02-04.
- [7] 杨义先,钮心忻. 安全通论(7):黑客篇之“战术研究”[EB/OL]. <http://blog.sciencenet.cn/blog-453322-956051.html>, 2016-02-14.
- [8] 杨义先,钮心忻. 安全通论(8):黑客篇之“战略研究”[EB/OL]. <http://blog.sciencenet.cn/blog-453322-958609.html>, 2016-02-25.
- [9] 杨义先,钮心忻. 安全通论(9):红客篇[EB/OL]. <http://blog.sciencenet.cn/blog-453322-960372.html>, 2016-03-04.
- [10] 杨义先,钮心忻. 安全通论(10):攻防一体的输赢次数极限[EB/OL]. <http://blog.sciencenet.cn/blog-453322-984644.html>, 2016-06-14.
- [11] 杨义先,钮心忻. 安全通论(11):信息论、博弈论与安全通论的融合[EB/OL]. <http://blog.sciencenet.cn/blog-453322-989745.html>, 2016-07-11.
- [12] 杨义先,钮心忻. 安全通论(12):对话的数学理论[EB/OL]. <http://blog.sciencenet.cn/blog-453322-993540.html>, 2016-07-30.
- [13] 杨义先,钮心忻. 安全通论(13):沙盘演练的最佳攻防对策计算[EB/OL]. <http://blog.sciencenet.cn/blog-453322-1000428.html>, 2016-09-02.
- [14] 杨义先,钮心忻. 安全通论(14):病毒式恶意代码的宏观行为分析[EB/OL]. <http://blog.sciencenet.cn/blog-453322-1001684.html>, 2016-09-08.
- [15] 杨义先,钮心忻. 安全通论(15):谣言动力学[EB/OL]. <http://blog.sciencenet.cn/blog-453322-1003586.html>, 2016-09-18.
- [16] 杨义先. 刷新你的安全观念,见杨义先的科学网实名博客[EB/OL]. <http://blog.sciencenet.cn/blog-453322-983276.html>, 2016-09-30.
- [17] 唐三一,肖燕妮. 单种群生物动力系统[M]. 北京:科学出版社,2008.