有组织的舆论风暴：商业网络攻击策略

Andrea Russo，意大利卡塔尼亚大学物理与天文学系，电子邮箱：andrea.russo@phd.unict.it

arXiv:2301.01518v1 [cs.CY] 4 Jan 2023

摘要——**良好的声誉对大多数组织和公司至关重要。事实上，良好的企业形象有助于企业与各类客户和合作伙伴建立更好的交易关系。然而，对于各种商业活动（企业对消费者、企业对企业、企业对企业对消费者、企业对政府）而言，声誉难以建立，却很容易毁于一旦。在与客户沟通的过程中产生误解，或仅仅是沟通策略不当，都可能给整个公司带来灾难。社交网络上数百万人的反应更凸显了这一点：如果他们对某一事件作出负面反应，就可能严重损害企业形象。这种现象被称为舆论风暴。**

本文提出一种针对组织发动舆论风暴攻击的周密策略，并展示对手如何利用这种攻击获取受攻击企业的私人信息。标准的企业安全程序并非为应对多域攻击而设计；因此，本文将展示如何通过实施不同类型的攻击，绕过传统且广为建议的安全程序。我还提出另一种舆论风暴攻击，能够有效地针对特定企业网络。最后，本文介绍了可减轻舆论风暴对公司造成负面影响的防御程序。

索引词—**舆论风暴、网络攻击、企业防御、社会动力学、压力测试、网络科学、《赛博朋克2077》**。

一、引言

在社交媒体出现之前，品牌危机在很大程度上是由记者的报道引发的。如今，舆论风暴是消费者数字口碑的集合，它指出了某种沟通失误，或公司犯下的某个严重错误 [15]。剑桥词典[1] 将舆论风暴定义为“突然发生、有时甚至激烈的反应”，将舆论灾难定义为“极其混乱且无法控制的局面、争议或一连串事件”。本文将交替使用这两个术语。

近年来，互联网上发生了许多舆论风暴 [19], [27], [31], ，这主要是由于社交网络用户数量不断增加。在某些情况下，人们曾正式研究舆论风暴，以更好地理解这一现象 [15], [28], [31]。2007年，几位研究人员就舆论风暴展开了讨论，其中一项主要结论是“自然科学模型适用于研究社会世界，但仍有一个核心问题尚待解决：社会世界能否、以及是否应当按照与自然科学相同的原则、程序和哲学进行研究” [1]。这一点很重要，因为如今我实际上能够利用来自数字世界的大量数据来研究和评估社会动力学，尤其是社交网络中的数据 [32]。

舆论风暴并非由单一事件构成，也不会遵循某种固定模式；相反，它们是由非线性动力学引发的，进而产生复杂的行为。因此，公司必须制定适当的程序，以应对各种危机情境。Lehtonen 理论 [23] 指出，舆论风暴会经历五个阶段：*(1)* 潜伏阶段，此时会收到危机即将发生的微弱信号；*(2)* 触发事件，此时相关对象成为新闻和社交媒体关注的焦点；*(3)* 相关对象登上头条新闻，媒体关注度飙升；*(4)* 媒体关注度逐渐平息，回落到一般性的哲学与伦理讨论水平；*(5)* 媒体仅有少量报道，关注点转向其他议题 [28]。

舆论风暴往往始于服务故障、社交失误或公司未能妥善沟通。因此，遵循适当的程序可以减少此类错误 [15]，。然而，现有的大多数质量与安全程序，例如 ISO 9001:2015 [17] 和 ISO/IEC 27002:2022 [18] 所建议的程序，都不足以应对多域网络与社交攻击。这是因为，就 27002:2022 而言，社交攻击不在其适用范围内；而 9001:2015 虽然着重提升业务流程质量，从而降低公众引发舆论风暴的风险，却无法缓解攻击者引发的舆论风暴。

因此，本文提出，攻击者可以蓄意发动舆论风暴攻击，以损害一家公司的声誉，并借此让竞争对手获利。我认为，自组织舆论风暴攻击需要大量已活跃于社交媒体的机器人：在这种情况下，机器人会针对目标公司掀起舆论风暴，传播假新闻（或放大某个事件，例如公司过去犯下的错误），引发大量真实用户作出负面反应，并在不知情的情况下继续替对手发动社交攻击。

此外，我认为，开源情报（OS- INT）可能让对手识别出组织的薄弱环节，也就是那些最有可能无法妥善应对舆论风暴或保护自己的人员，因而无法及时减轻其影响。许多员工都拥有 LinkedIn、Facebook 或 Twitter 账户：将舆论风暴引向目标公司员工的社交媒体账户，可能会让

[1]https://dictionary.cambridge.org

员工承受极大的压力。对于那些不常处理公共关系的人来说，情况可能更糟，并可能引发困惑、恐慌和苦恼。事实上，舆论风暴发生时，即使是负责沟通过程的人员和管理者也可能惊慌失措，而担心失去客户和合作伙伴可能会对任何公司造成极为不利的影响。

当目标企业的员工处于这种状态改变之中时，我认为可以制定一种社会工程策略来获取受保护信息：在这种情况下，舆论风暴不仅用于损害企业形象，也被用作社会工程攻击的掩护。事实上，尽管大多数重要组织都遵循ISO/IEC 27002 等安全标准中列出的最佳实践：2022 [18], 在舆论风暴这类社交攻击期间，由于负责遵守相关程序的人员承受压力，某些最佳实践和程序可能会被有意或无意地曲解或绕过[14]。

贡献。本文的贡献如下：
1) 我说明了如何发起自动化且有组织的 Firestorm 攻击，只需进行少量手动操作，例如选择主题和话题标签；2) 我介绍了攻击者在实施舆论风暴时可能采取的操作分类体系；3) 我阐述了舆论风暴的发起者如何通过针对个别员工而非公司账号来发动攻击、逃避检测，同时加大对企业造成的损害。4) 我展示了公司可采取哪些短期和长期程序来减轻舆论风暴攻击的影响。

II. 网络攻击计划前奏

在本节中，我将介绍一种新颖的策略：利用僵尸网络煽动真实用户反对目标公司，从而人为制造舆论风暴。机器人能在数秒内发布大量帖子，因此可用于在社交网络上放大任何观点，影响政治事务 [3]和企业价值 [33]。例如，某报纸的 Twitter 账号遭到网络攻击后，发布了一则关于奥巴马总统在白宫被炸弹炸伤的假新闻，引发华尔街闪崩，并导致所有经济交易暂停数分钟。这使标准普尔500指数及其相关公司损失约1210亿美元 [11]。

我将攻击计划分为六个阶段：
1) **寻找可作为 Firestorm 攻击基础的事件/话题**。 这可以是过去发生的事件，也可以是企业过去犯下的错误，并以此作为即将展开的攻击的依据。我将这一事件定义为目标话题。
2) **利用机器人制造或放大潜伏状态**。通过利用僵尸网络，对手可以在社交媒体上发布大量帖子，让目标话题触及更多人，并使他们有机会作出负面反应。这最终可能导致一种状态真实用户开始自发谈论这一主题，并开始自行传播有关目标话题的信息。为此，攻击者可以重复使用旧的热门话题标签，或创建新的话题标签：话题标签是一个关键词，其所代表的信息会促使人们采取社交行动。
3) **让话题在人群中传播**。对攻击者而言，理想的情况是，真实用户看到僵尸网络发布的帖子后，开始发布与目标话题相关的内容。这会让话题受到更多关注，并可能使其成为热门话题。例如，Twitter 允许用户查看当前流行的话题和话题标签。如果这种情况发生，就会出现这样一个时刻：有足够多的人发布与目标话题相关的内容，使舆论风暴能够持续数天，而攻击者的僵尸网络无需再发布任何帖子。我把这一时刻称为燃点。[2] 相反，如果真实用户没有对该话题作出负面反应，或者话题触及的人数不足以让舆论风暴达到燃点，围绕该话题的讨论就会逐渐降温，最终结束。在这种情况下，我称舆论风暴熄灭了。不过，攻击者可以更换目标话题，并从第1阶段重新开始。

4) **识别人类目标**。管理者（例如，首席技术官、首席执行官）是公司的决策者。攻击者可能会想要列出这些人的名单，以便在攻击从公司的社交网络个人资料转移到员工的个人资料时使用这些姓名。找出那些以在受攻击的公司工作为荣的人，也有助于对公司施加更大压力（因为他们与公司的价值关系更密切）。

5) **将焦点转向员工**。在舆论风暴最为活跃期间，先前制造潜伏状态的那些机器人会将注意力转向遭受攻击的企业员工所拥有的公开社交媒体个人资料。这些个人资料已在攻击的上一步中被识别出来。这可能会使舆论风暴的关注焦点转向员工，也让他们感到不适。由于品牌通常是舆论风暴的中心，聚焦于员工会对他们造成更大影响，也可能扰乱内部流程。

6) **实施网络攻击**。由于人们会减少对遵循内部规程的关注，公司采取的许多安全最佳实践可能无法得到妥善遵守，甚至可能被忽视。攻击者可以利用这种行为为自己谋利。

为了将关注焦点从公司转移到员工身上，有必要优化这一转变的时间尺度和时机，因为人们攻击员工并非一个线性过程，但如果负面事件造成的负面影响和价值都很高，这种攻击就更容易发生。将攻击目标转向员工

[2]在化学中，燃点 是某种燃料被点燃后能够持续燃烧至少五秒钟的最低温度。

还有另一种副作用，这对攻击者有利：各国负责公共网络安全的组织无法在公司页面上看到针对公司的 Firestorm 攻击，因为舆论风暴只针对员工。这些组织很难发现所有针对员工的评论和帖子，从而让攻击者制造出一种烟幕式攻击，绕过传统的安全措施、程序和策略。由于这些组织必须主要关注遭受攻击的公司，因此可能不会投入太多精力，分析针对该公司所有运营人员的每一次互动。

III. 商业社交情绪疾病与网络策略

剑桥分析事件凸显了社交媒体对大多数人群和组织的作用与重要性。美国司法部为调查外国可能对美国施加的影响而编写的一份文件，展示了确实存在一些旨在影响个人及公共和私营组织的机构（例如 IRA——互联网研究机构） [36] [29]。

要成功影响人们，很大一部分关键在于了解系统的初始条件，也就是通过社交网络上获取的数据，对这些人进行准确的画像分析。对某些议题较为敏感的人，以及能够对其生活和工作所在社群产生最大影响的关键人物，是社交攻击的主要目标，因为他们在网络中处于中心位置（枢纽）。

画像分析是指通过数据收集及后续处理，全面或近乎全面地了解一群个人或某个个人，掌握他们的习惯和偏好 [13]。所获取的信息涉及政治、音乐和社交兴趣，包括识别他们的朋友网络、同事等。借助这些信息，传递任何内容都会容易得多，因为我们可以了解哪些人最容易受到各种话题影响、对哪些话题感兴趣，并触及他们的弱点、恐惧和兴趣。此外，还可以推断谁可能通过自己的网络传播某项内容；如果相关对象是具有重要或核心角色的人，成功的可能性就会呈指数级提高。

剑桥分析在 2016 [36] 的全国选举期间，使用与人格特质相关的 OCEAN 模型，了解美国许多人的偏好。OCEAN 模型可以向对某个主题敏感的人发送特定信息和内容。这种方法与传统、标准的大众传播方式截然不同，因为它能够将合适的内容发送给合适的人。不幸的是，剑桥分析丑闻被定性为老式的传统政治影响手段，包括卖淫、偏袒等。事实上，正如 Brittany Kaiser（剑桥分析前业务发展总监）在 2018 年下议院文化委员会的质询环节中所说的那样，这起丑闻揭示了“一种新型武器”，用来描述剑桥分析所做的工作，也将人工智能视为真正的软实力武器 [13]。

然而，仅仅了解员工关注的热点还不够——要改变他们的情绪并发起有效的社交攻击，还需要找到一个主题。在社交网络上，舆论风暴期间，人们通常会被三类失误所触发 [15]:

1) 社交失误 2) 沟通失当
3) 产品或服务失当

虽然这些事件看起来相似，但不同类型的事件可能引发不同的动态和反应。例如，在产品或服务故障的情况下，绩效相关危机会让人们质疑品牌是否有能力实现基本的功能性表现 [9]。另一项研究还发现，舆论风暴不仅会对品牌造成短期影响，也会产生长期影响；该研究至少追踪了最近一次舆论风暴发生两年后的情况 [15]。

下面分别举例说明上述每一种触发因素。

1) **社交失误**。企业可能是某起事故或犯罪的共犯，例如 Nike 的童鞋事件 [10][30]，或 2012 [31]中的 ING-DiBa 案例。2) **沟通失当**。企业可能未能妥善沟通，例如发表针对某个社群或运动的负面评论 [27]。3) **产品或服务失当**。企业可能销售会伤害消费者的产品，例如可能致人死亡的疫苗 [19]。

这些失误及其引发的舆论风暴，可能会让受影响的员工感到不安和恐慌，因为同事、朋友以及其社交网络中的其他人可能会将受影响的员工视为舆论风暴的根源。

社交网络攻击还会给公司带来意想不到的负面后果：

1）公司在金融市场上的价值可能会迅速下跌； [11]2）舆论风暴期间在公司工作的人，未来可能遭受歧视，尤其是当舆论风暴由本可避免的（据称）不可接受的失误引起时 [26], [38]。3）与员工一样，受到冒犯的品牌也可能背负长期污名，促使其他公司向遭攻击企业的员工发出工作邀约。这会使该企业处于更加不利的境地，因为员工会受到激励，离开遭攻击的公司并接受新的工作邀约。

网络以及新闻的重要性和范围，都可能对公司的反应和动态产生重大影响。网络以及新闻的重要性和范围，都可能对公司的反应和动态产生重大影响。例如，当一家公司的

员工收到一则重要性很高的新闻时，可能会因新闻的重要性而感到无能为力；他们觉得自己无需承担责任，因为事件的影响超出了自身行动所能及的范围，因此往往会将大部分责任推给公司的管理者。

确实，在混乱或无序时期，秩序越少，《熵》就越高；秩序越多，紧急程度就越高：这是因为组织内的人意识到了紧急情况，组织也会自我完善以应对这一情况 [39]。

当公司里的许多员工陷入恐慌时，组织的首席传播官（CCO）会在公司页面上详细说明并应对舆论风暴；然而，这无法阻止针对员工个人主页的社交攻击。因此，即使是负责管理沟通过程的人员和管理者也可能陷入恐慌，因为舆论风暴持续的时间越长，失去客户和声誉的可能性就越大。对任何公司而言，这都是一种糟糕的处境，尤其是在多年努力之后。然而，管理者在组织架构图中被视为“关键员工”，因此，考虑到他们在公司承担的职责，他们不应受到社交操纵和社交疾病的影响。上个世纪，这类组织架构图通常呈金字塔形，首席执行官位于顶端；如今，敏捷模型则允许公司以不同方式安排组织架构图中的员工。不过，每个领域的错误或重大问题，其法律和个人责任始终由该领域的最高管理者承担——例如，网络安全通常由首席信息安全官（CISO）负责。网络侧策略能够强烈影响接近管理者和主管的员工，直接影响团队的情绪，包括管理者本人。更具体地说，攻击者会利用公司的网络枢纽，也使公司的其他员工失去防备。一旦社交疾病已在公司内部广泛传播，许多人又因舆论风暴而承受压力，网络攻击就可能开始。

IV. 评估攻击面

本节将介绍对手（或参与舆论风暴的真实用户）可能采取的操作，这些操作可能进一步扰乱目标公司的业务流程、损害其企业形象，或获取机密信息。为此，我将对这些操作进行一种新的分类，并分析它们对信息安全基本属性——机密性、完整性和可用性 [34]——的影响。

我将这些操作分为三类：

1) **控制大规模实体**，即数千甚至数百万个不同的参与者同时对一家企业采取多种操作。这些参与者既可以远程行动，也可以亲自行动；既可以是机器人，也可以是人类。

2) **利用内部人员**，即利用员工所犯的错（例如，他们可能因为舆论风暴而承受压力），或安排一名内部威胁者提取机密信息。

3) **索要赎金**，即对手可能会要求支付赎金，以停止舆论风暴。这可能导致机器人被关闭，甚至转而在社交媒体上为公司辩护。

我将在此分析每个类别中的不同操作及其影响。分析结果总结于表 I 中。

A. 控制大规模实体

a) 拒绝服务（*DoS*）攻击：对手可能想通过破坏其所提供服务的可用性来损害企业的声誉。为此，攻击者可以利用僵尸网络，每秒向目标服务发送大量请求，使服务器不堪重负，导致服务中断。如果可能，攻击者甚至可以重新利用用于制造潜伏状态的僵尸网络，并为其重新部署拒绝服务脚本。或者，如果对手并非单一实体，而是一大群有组织的人，他们也可以通过简单脚本发动拒绝服务攻击，无须利用任何僵尸网络，因为大量对手能够产生足以使服务器过载的流量。不过，在这种情况下，对手必须谨慎安排攻击时机，并可能希望隐藏自己的位置，例如使用虚拟专用网络。最后，对手还可以鼓动真实用户使目标企业的服务器过载，因为他们可以利用用于话题标签宣传的机器人账号来协调攻击。

b) 物理行动：业务流程也可能因合法但有害的物理行动而中断或放缓。例如，有人在企业办公场所周边示威：示威者在大楼外聚集，堵塞通往办公场所的道路或拖慢通行速度，或导致前往大楼的交通比平时更加拥堵，使员工无法按时到达工作地点。另一个例子是有人打电话到组织的呼叫中心，唯一目的就是抗议。

B. 利用内部人员

a) 人为错误：尽管众所周知，人为错误是导致安全事件的最主要原因之一 [16], [43], ，但大多数公司仍未充分投入资金为员工提供培训，导致数据泄露或其他安全相关事件发生 [22]。这意味着，如果攻击者想要初步侵入目标组织的系统，根据员工识别钓鱼邮件或诈骗网站的能力，他们或许无需发动舆论风暴攻击也能得逞。然而，遭遇舆论风暴的员工，无论舆论风暴针对的是他们任职的公司还是他们自己的个人资料，都会因感受到危机而更倾向于违反内部政策，从而犯下错误 [2]。

b) 提供帮助：在舆论风暴最为猛烈时，对手会主动联系遭受攻击的企业，假扮成能够帮助其减轻舆论风暴影响的专业人士（例如顾问），比如声称自己曾以社交媒体经理的身份处理过舆论风暴。这种情况可能

通过电子邮件、社交网络或企业网站发生，例如当企业有职位空缺时，对手假扮成求职者。对于规模较小的企业，对手甚至可能亲自前往遭受攻击的公司所在地。如果攻击者成功受聘，便可能接触到机密信息。我认为，攻击者不会想在自己还是公司员工期间篡改文件或攻击公司基础设施。

c) 内部人员威胁：对手可能不会亲自加入企业，而是联系仍在遭受攻击的公司任职、但未在社交媒体上表示支持，甚至对公司表达不满的员工。攻击者可能会试图说服他们分享机密信息，使他们成为内部人员威胁 [25] ——如果得逞，攻击者不仅能获取机密信息，如果窃取的内容还会损害企业利益，也可以将其发布到网上，进一步损害企业的声誉。

C. 索要赎金

a) 勒索以停止攻击：对手联系遭受攻击的企业，并证明发动舆论风暴的僵尸网络由自己控制。随后，他们要求以比特币支付任意金额，以关闭机器人，从而（希望能）阻止攻击的大部分。事实上，如果舆论风暴已经成功煽动许多人加入社交攻击，那么关闭僵尸网络可能无法阻止或减缓舆论风暴。如果对手计划用舆论风暴攻击多家企业，就会避免此类情况，因为受害者支付赎金的可能性，与攻击者收到钱后停止攻击的可信度成正比。换句话说，只有当攻击者被认为会在收到赎金后停止攻击时，受害者才更有动力付款 [4]。

b) 防御即服务：对手联系遭受攻击的企业，但他们并不表明自己是攻击的幕后操纵者，也不索要钱财来停止攻击，而是试图向受害者兜售舆论风暴防御服务，声称该服务由机器人组成，能够维护企业声誉：这基本上是一场反向舆论风暴，那些先前营造潜伏状态的机器人如今转而保护公司。为了避免引起过多关注，攻击者可能会缓慢调整攻击机器人与防御机器人之间的比例，直到所有机器人都转而保护公司。

V. 案例研究：CD PROJEKT RED

2020年12月10日，CD PROJEKT RED发布了备受期待的游戏《赛博朋克2077》。该游戏在发售前就已广受欢迎，在开发过程中持续引发电子游戏社区的热议，并连续两年获得金摇杆奖“最受期待游戏”奖项。 [42]如图1和图2所示，在游戏发售前的10天里，热度大幅攀升，并于12月10日达到顶峰；当天，Twitter 上的话题标签 # Cyberpunk2077 被推文提及了193,900次，

表 I 社交攻击面评估

类别	操作	影响			
		机密性	完整性	可用性	Rep.
大规模	拒绝服务攻击	No	No	Yes	Yes
	物理动作	No	No	Yes	Yes
内部人员	人为错误	Yes	Yes	Yes	Yes
	提供帮助	Yes	No	No	No
	内部威胁	Yes	No	No	Yes
赎金	敲诈	No	No	No	No
	国防服务	No	No	No	No

机密性： 该操作可能影响机密性。 | **完整性：** 该操作可能影响完整性。 | **可用性：** 该操作可能影响可用性。 | **声誉：** 该操作可能对公司的声誉造成负面影响。

这些推文来自53个不同国籍的用户。在这段时间里，许多其他与该游戏相关的话题标签也非常热门，例如 # Cyberpunk2077Hype 被转推了10,000次 [41]。

然而，游戏发布几天后，《赛博朋克2077》的话题再次兴起，这次与补丁和退款相关的讨论有关。事实上，游戏发布得太早，存在许多漏洞：因此，有几位玩家向 CD PROJEKT RED 申请退款，而且常常还会在在线商店给游戏留下差评。这在公司内部造成了一种“信息疾病”，正如第三节所述：在本例中，CD PROJEKT RED 的员工感到压力重重，并因《赛博朋克2077》的品质而倍感压力；他们为这款游戏投入了两年多的辛勤工作。 [42]

2021 年二月初，也就是游戏发布仅 60 天后，CD PROJECT RED 遭到勒索软件攻击，攻击者成功窃取了多款游戏的源代码，包括管理文件 [8]。随后，攻击者威胁公司，除非企业向网络犯罪分子支付巨额款项，否则就泄露或出售窃取的代码和文件。最终，CD PROJECT RED 拒绝与攻击者谈判，并在新闻稿中表示，他们“不会屈服于要求，也不会与该行为者谈判”；同时确认此次攻击未获取任何个人信息，并表示正与执法部门合作追查攻击者 [7], [35]。之后，安全分析师发现，被盗源代码正在暗网上拍卖，起拍价为 100 万美元。 [40] 攻击者表示已收到令其满意的报价后，拍卖便告结束 [40]。在这场拍卖结束后一周内，代码通过社交媒体在网上传播，CD PROJECT RED 开始使用 DMCA 下架通知移除包含其代码的帖子 [24]。

CD PROJEKT RED 为《赛博朋克2077》营造的社交媒体热度，被黑客利用来威胁公司并勒索钱财；但这也产生了副作用，即损害公司的声誉，进而可能影响其他备受期待的游戏的销售额。

表 II 展示了情感倾向分

析结果，这些结果来自话题标签 #CDprojectRED 下的推文和评论。收集自 Twitter 的数据遵循《赛博朋克2077》的发行及开发时间线；表中数据可分为三个类别：发行前（十月和十一月）、发行期间（十二月和一月）以及《赛博朋克2077》发行后（二月）。

可以看出，十月和十一月的情感倾向总体保持中性偏正，期间略有波动。十二月游戏发布时，由于游戏中存在大量漏洞，负面情感倾向略有上升，但这一增幅相当有限。一月，随着更多玩家开始玩这款游戏，负面情感倾向变得强于正面情感倾向，不仅使复合情感得分为负值（-0.111），也使玩家对游戏和开发者的情感倾向转为中性偏负。最后，到了二月，整体情感倾向恢复中性，但与十月和十一月相比，负面情感倾向仍然更强。

这些数据表明，CD PROJEKT RED 公司在游戏发布期间承受了多大的压力。此外，图3展示了整个游戏发布期间公司的财务价值，并标出了发生的两个关键事件：黄线表示游戏发布，红线表示勒索软件攻击。可以看出，游戏发布后，公司的财务价值骤然下跌；这很可能是因为游戏中存在许多漏洞、评价不佳且受到批评，导致客户对公司失去信任。接下来的两个月里，公司收复了超过一半的损失；然而，勒索软件攻击又导致公司财务价值下跌，因为客户再次对公司失去信任，这次是出于安全方面的考虑。

表 II Twitter 上 #CYBERPUNK2077 的 VADER 情感倾向

月份	负面	中性	正面	复合情感得分
十月	0,085	0,757	0,150	0,163
十一月	0,079	0,766	0,149	0,163
十二月	0,087	0,750	0,161	0,153
一月	0,143	0,758	0,093	-0,111
二月	0,104	0,745	0,145	0,120

VI. 商业防御策略

为避免公司遭遇危险事件，人为因素至关重要 [37]，，但也可以制定具体的防御策略。第三节介绍的失误，即社交失误、沟通失败以及产品或服务故障，都可以加以分析，以防止事件发生。对我们大多数人来说，某条信息（例如迷因、话题标签）在短时间内“走红”、触达数百万个节点的消息，似乎纯属偶然，因而无法预测；但 Kolli 等人 [21] 发现，至少在 20% 的情况下，级联规模的变化看起来是随机的，而在其余 80% 的情况下，则有可能

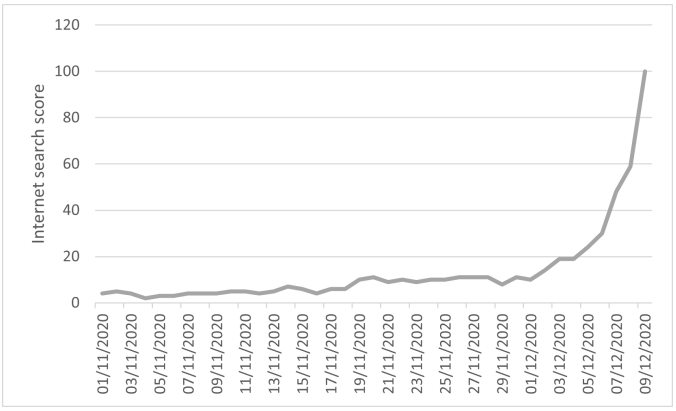


图 1.《赛博朋克2077》发售的社交热度兴趣分数

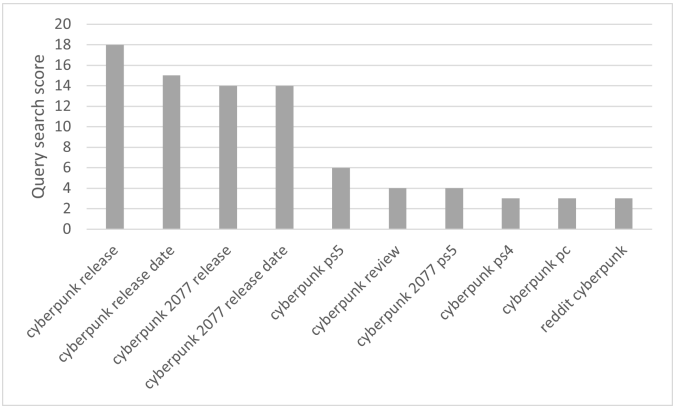


图 2.《赛博朋克2077》发售的社交热度搜索量

预测级联规模未来的变化。因此，可以制定短期策略，在 Firestorm 攻击尚处于早期阶段、潜伏状态正在形成时及时发现它们。不过，也可以通过主动治理制定长期防御策略。以下是一种可能的长期主动策略：

1) 制定公司内部流程，帮助员工保护自己免受社交媒体（如 LinkedIn）上的各种攻击；2) 制定公司外部流程，例如联

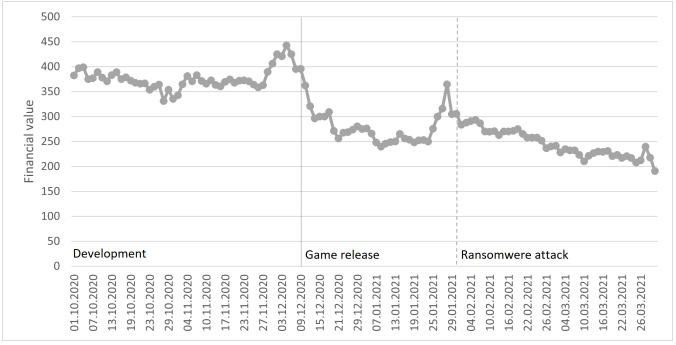


图 3. CD PROJEKTRED 的财务价值与关键事件

系盟友／合作伙伴公司，请求他们协助应对社交媒体上的各种攻击；3) 预先创建辅助机器人，以自动保护公司；4) 建立一个记录发起过舆论风暴的账户的国际数据库。所有公共和私人机构均可访问该数据库，这将有助于判断正在发生的舆论风暴是真实的还是人为制造的。[12]

大众媒体可以凸显这三种可能的操作，并公开表明舆论风暴之所以得到应对，是因为其他个人或组织开始为遭受攻击的公司辩护。因此，与无所作为相比，这些操作能让舆论风暴更快平息，并最终消散。[15] 如果一家公司过去犯下了极其严重的错误，那么每当这家公司再次犯错时，都有可能再次引发舆论风暴，针对近期事件，也可能再次针对过去的事件。事实上，舆论风暴可能会在约 2 年后卷土重来 [15]。

发生社交失误时，还会产生一种需要缓解的额外副作用：舆论风暴会自然而然地蔓延至员工，而无需对手操纵。针对这一副作用的示例防御策略可按以下方式实施：

1) 让公司内外的人在社交网络上就相关话题展开对话（例如 ING-DiBa 的肉食者与素食者之争 [31]）。这一策略可以增加关注者数量；2) 将公司以外的某个实体归咎为替罪羊，使舆论风暴从公司转移到指定实体。虽然这不太道德，但通常确实有效；3) 根据攻击的强度、持续时间和波及范围，可以为公司制定应对策略。

a) 社交失误：如果舆论风暴与合作公司有关，或公司只有某个特定部门遭到攻击，应立即与其划清界限。

b) 沟通失当：这里的目标是维护公司的声誉和权威。在这种情况下，应立即与沟通失误撇清关系，并继续执行公司的声誉策略，让事情看起来只是一次意外。此外，为此事道歉也无妨。

c) 产品或服务失当：立即停止生产受影响的产品或提供相关服务。成立一个委员会，评估产品/服务的质量。尽管考虑到合作伙伴的数量、质量标准和企业连续性，这项操作可能较为复杂，但如果及时采取，就能在沟通层面形成良好的防御屏障，因为人们会明白公司本身也已认识到问题，从而限制损害；

时机在舆论风暴期间至关重要。首先，要判断舆论风暴是真实的还是人为制造的（可以查看发起舆论风暴的账户创建日期来判断——如果最初的账户是最近才注册的，它们很可能是机器人，因此这场舆论风暴是人为制造的）；其次，要加强网络防御，为可能发生的网络攻击做好准备；第三，要考虑公众反应，因为这意味着受影响的公司比其他人（在社交网络上制造舆论风暴的人）更早或同样及时地发现了问题，并会迅速采取行动解决问题，让客户放心，问题将得到解决。这有助于平息或消除舆论风暴。例如，ING-DiBa 的肉食者与素食者之争就是由沟通失误引起的。此前，该公司 Facebook 专页从未有过如此庞大的流量，他们认为这是增加关注者数量的机会。事实上，舆论风暴发生几天后，攻击者仍在发帖，新获得的关注者也加入了争论，开始为公司辩护。[31]

显然，公司有必要根据舆论风暴的类型（真实或人为）调整策略，以应对相应类型的攻击（真实或人为）。当然，预防措施在这两种情况下都适用，但了解自己正在与谁对抗以及事态的成因，有助于维护公司的声誉，有时甚至能保住公司本身。

VII. 未来工作

在接下来的研究中，我希望采用不同的压力动态，也就是说，既可以实施迅速、大规模且猛烈的舆论风暴，也可以实施持续不断、涉及账户较少的舆论风暴。根据舆论风暴的类型，这些动态会以不同方式影响公司和员工，或许可以表明，对某些公司而言，持续性的舆论风暴更有效，而对另一些公司而言，迅速爆发的舆论风暴更合适。未来工作中我还想探讨的另一个方面是如何联系公司内部人员，也就是说，采用更可能引发道德反应的信息。例如，机器人联系员工，并指出其给公司造成的损失。这种情况非常有意思，因为在对员工进行“道德劝诫”之后，就可以运用社会工程策略来促成网络攻击。另一方面，对于公司外部、不以员工为目标的情形，则可以采用一些策略来提高网络攻击得手的概率，或实施信息或金钱勒索。例如，在舆论风暴期间，可以联系遭受攻击的公司，冒充国家网络安全机构，并采取如下策略：

1) 冒充国家网络安全机构，声称大多数账户都是虚假账户，并套取有关其安全防护的信息；

2) 冒充国家网络安全机构，进入其计算机系统。3) 冒充国家网络安全机构，声称正在发动网络攻击以测试其网络防御，随后立即发动第二次攻击，利用第一次攻击获取的信息并泄露部分防御措施；或者声称自己没有抵御第一次攻击，以获取所需数据。

无论如何，这类互动都将通过计算机模拟来开展，因为出于显而易见的伦理原因，实施这些策略是不可能的，至少也是极其困难的。

VIII. 结论

本文阐述了与网络安全相关的一些事件如何与特定的社会动力学相联系。当社会动力学相互交织并与网络目标相结合时，便无法再将攻击类型简单归类为传统的攻击（网络攻击或社交攻击），而应视为社会网络攻击，因为其中一种攻击的有效性也会提高另一种攻击成功的可能性。我提出了一种新模型，使研究人员和公司能够：（1）了解公司和组织何时容易受到社会网络攻击；（2）说明公司和组织如何保护自己免受舆论风暴影响；（3）证明社会网络攻击应被界定为跨多个领域可能造成高风险的事件；（4）展示一种新的网络攻击模型，并采用多学科社会学方法，提升常见网络攻击的潜在影响。基于 CD Projekt Red 事件案例收集的数据表明，尽管人们对这类攻击仍知之甚少，但它们未来可能会成为常态，因为公司的资产不仅包括人力资本以及商品和／或服务的生产，也包括公司自身的声誉。

IX. 作者及论文信息

A. 数据收集

我使用 Tweepy 和 Twitter档案API 收集与 # Cyberpunk2077 相关的推文。这两种服务都获得了 Twitter 的许可来获取和收集数据，但任何下载的主题数据都需要经过修订和清理，以提高研究质量。例如，我发现许多复制粘贴的推文（由垃圾信息行为或虚假账号/机器人造成），还有几条推文在进行 VADER情感分析时包含 VADER程序无法识别的词语，因此我将其删除。对于每个主题，我都使用相同的方法论来获取标准且高质量的数据。此外，为了获取每天/每小时的准确推文数量（定义为推文条数），我使用 getdaytrends.com；这是一个专门的网站，可以实时监测各个主题，也可以查看过往主题。总计，我们的数据超过~5000条推文。我从 https://www.investing.com/equities/cdproject-historical-data 网站获取了 CD Projekt Red 的财务数据。

B. 作者贡献

调查与数据资源、方法论、数据清理和软件：A.R.；所有作者均已阅读并同意发表此稿件版本。

C. 资金支持

作者披露了以下用于本研究、论文撰写和/或本文发表的资金支持：本项目获得了卡塔尼亚大学的资助。

D. 作者简介

Andrea Russo 是卡塔尼亚大学复杂系统专业的博士候选人。目前，他在物理与天文学系工作。他曾与 CNR Ibam 合作，也曾专门参与技术与社会相关项目。

他的主要研究领域和兴趣集中于计算社会方法的研究与发展，旨在解释社会复杂性，尤其关注政治、经济学、商业和国防安全领域的应用。

Orchid：0000-0003-3816-0539 通讯作者。电子邮箱：Andrea.russo@phd.unict.it

我想感谢“Vereos”和“Andrea metal clone”，他们帮助我构思并完善了这篇论文。

参考文献

[1] E. B. Alan Bryman.《商业研究方法》，第2版。牛津：牛津大学出版社。牛津大学出版社，2007年。[2] L. Bakos、D. D. Dumitraș cu 和 K. Harangus。信息物理系统去中心化危机管理与沟通中的人为因素准备。《可持续性》，11(23):6676，2019年。[3] G. Carrer 和 F. Bechis。《中国如何利用机器人在意大利开展宣传：Alkemy 为 Formiche 所做的 Twitter 分析》。*Formichiere.it*，第1页，2020年。[4] E. Cartwright、J. Hernandez Castro 和 A. Cartwright。付还是不付：勒索软件的博弈论模型。《网络安全杂志》，5(1):tyz009，2019年。[5] M. C. Ciccarelli。丑闻发生后重建员工信任。《人力资源主管》，2018年。[6] K. Creighton。公司遭遇公众丑闻后，如何恢复员工信任。《hrdailyadvisor》，第1页，2019年。[7] C. Criddle。《赛博朋克 2077》开发商 CD Projekt 遭勒索软件攻击。*bbc.com*，2021年。[8] D. D. CD Projekt 遭黑客攻击，《昆特牌》源代码泄露。*eip.gg*，2021年。[9] N. Dawar 和 M. M. Pillutla。产品伤害危机对品牌资产的影响：消费者期望的调节作用。《营销研究期刊》，37(2):215–226，2000年。[10] J. Day。Nike：“无法保证没有童工”。《卫报》，2001年。[11] M. Farrell。高速交易引发 Twitter 闪崩。*CNN Business*，2013年。[12] FrancescoArruzzoli。《网络威胁情报在组织中的作用》——Zoom。[13] G. Giovanni 和 A. Russo。社会画像与国家安全。*SOCINT Press*，2021年。[14] G. Halkos 和 D. Bousinakis。压力与满意度对生产率的影响。《国际生产率与绩效管理期刊》，2010年。[15] N. Hansen、A.-K. Kupfer 和 T. Hennig-Thurau。数字时代的品牌危机：社交媒体风暴对消费者和品牌的短期与长期影响。《国际营销研究期刊》，35(4):557–574，2018年。

[16] K. Hughes-Lartey、M. Li、F. E. Botchey 和 Z. Qin。人为因素是组织物联网信息安全中的关键薄弱环节。Heliyon，7(3):e06522，2021年。[17] iso.org。ISO 9001:2015。iso.org，2015年。[18] iso.org。ISO/IEC 27002:2022。iso.org，2022年。[19] A. G. Kate Connolly 和 J. Henley。牛津疫苗暂停使用后，德国和意大利陷入混乱。《卫报》，2021年。[20] R. Knight。如果你的公司正陷入公众丑闻，你该离开吗?《哈佛商业评论》，第1页，2018年。[21] N. Kolli、N. Balakrishnan 和 K. Ramakrishnan。利用熵量化在线社交媒体级联的可预测性。载于《2017年IEEE/ACM社交网络分析与挖掘进展国际会议论文集》，第109–114页，2017年。[22] P. Langlois。《2020年数据泄露调查报告》，2020年。[23] J. Lehtonen。《危机沟通》。芬兰广告商协会，1999年。[24] Lorenzo。CD Projekt Red 运用《数字千年版权法》下架分享被盗游戏代码的推文，2022年。[25] G. Mazzarolo 和 A. D. Jurcut。网络安全中的内部威胁：来自内部的敌人，2019年。[26] K. McLeod。HES 丑闻后陷入困境的员工称，老板有钱支付工资，却拒绝支付。Daily Record，第1页，2019年。[27] M. Monkey。Twitter 用户不爱麦当劳了。《卫报》，2012年。[28] K. Nuortimo、E. Karvonen 和 J. Härkönen。通过大型数据集媒体分析建立社交媒体风暴量表。《营销分析期刊》，第1–10页，2020年。[29] 美国司法部。关于调查俄罗斯干预2016年总统选举的报告。美国司法部，2019年。[30] J. Oliver。从布伦特·斯帕尔号事件中吸取教训。Politico，1995年。[31] J. Pfeffer、T. Zorbach 和 K. M. Carley。理解在线舆论风暴：社交媒体网络中的负面口碑传播动态。《营销传播期刊》，20(1-2):117–128，2014年。[32] F. M. Rinaldi、G. Giuffrida 和 T. Negrete。实时监测与评估——将新兴新闻作为预测流程：一种基于大数据的方法，2017年。[33] R. R. Riverso。巴塞罗那，前主席巴托梅乌被捕，2021年3月。[34] S. Samonas 和 D. Coss。中央情报局反击：重新定义安全领域的机密性、完整性与可用性。《信息系统安全期刊》，10(3)，2014年。[35] J. Schreier。CD Projekt 勒索软件攻击严重干扰《赛博朋克》更新工作。bloomberg.com，2021年。[36] 美国参议院。关于“评估俄罗斯对近期美国选举的活动与意图”的背景：分析流程与网络事件归因。美国参议院，2017年。[37] C. Simonelli。先教育，再购买：人为因素在打击勒索软件中的作用。Formiche.net，第1页，2021年5月29日。[38] K. Strauss。大众汽车如何在排放丑闻后鼓舞员工（至少目前如此）。福布斯，第1页，2017年。[39] M. Tang 和 X. Mao。用于衡量人工社会中涌现现象的基于信息熵的指标。《熵》，16(8):4583–4602，2014年。[40] Topic。报道称 CD Projekt Red 源代码在暗网拍卖中以数百万售出 [更新] | Ars Technica，2022年。[41] A. User。#Cyberpunk2077Hype • 美国 • Twitter热门话题标签，2022年。[42] 维基百科贡献者。《赛博朋克2077》。https://it.wikipedia.org/w/ index.php?title=Cyberpunk_2077&oldid=130410919。访问日期：未知-未知-未知。[43] C. C. Wood 和 W. W. Banks。人为错误：一个被忽视但影响重大的信息安全问题。《计算机安全》，12(1):51–60，1993年2月。