

入侵分析钻石模型

Sergio Caltagirone
sergio.caltagirone@cciatr.org

Andrew Pendergast
andrew.pendergast@cciatr.org

Christopher Betz
christopher.betz@cciatr.org

“情报分析人员应该对他们的推理过程保持自我意识。他们应该思考如何做出判断并得出结论，而不仅仅是依靠判断和结论本身。”——Richards J. Heuer Jr. [1]

“tcpdump对于入侵分析来说就像望远镜对于天文学一样重要。”——Chris Sanders [2]

摘要

本文提出了一种由分析人员构建的入侵分析新模型，这个新的模型是源于分析人员在多年的分析工作中提出的一个简单问题：“对我们工作来说什么是基本的方法”。这个模型建立了任何入侵活动的基本原子元素，即事件，事件由四个核心特征组成：对手、基础设施、能力以及受害者，这些特征是相互关联的，表示了它们之间的基本关系，又恰好形成了钻石的形状，因此我们给这个模型取名为钻石模型。该模型进一步定义了额外的元特征来支持更高级别的构造，例如将事件链接到活动线中，并进一步将事件和活动线融合到活动组中。这些元素、事件、活动线和组都有助于围绕分析过程构建入侵活动的基本而全面模型。钻石模型定义了入侵分析和对手行动的基本概念，同时允许模型灵活扩展并包含新的思想和概念，该模型首次建立了一种将科学原理应用于入侵分析的形式方法，特别是在可测量、可校验和可重复方面提出了一种综合的活动记录，活动组合和活动关联分析方法。这种科学的方法极大地提高了分析的有效性，效率和准确性。最后，该模型能够实时结合人工情报实现网络防御，跨事件自动化关联，将事件分类到对手活动中，并在规划和执行缓解策略的同时预测对手行动。

目录

1	引言	5
2	相关工作	6
3	钻石模型概述	7
4	钻石事件	8
4.1	对手	11
4.2	能力	12
4.2.1	命令与控制	13
4.3	基础设施	13
4.4	受害者	14
4.4.1	系统漏洞	14
4.5	事件元特征	15
4.5.1	攻击时间戳	15
4.5.2	攻击阶段	15
4.5.3	攻击结果	16
4.5.4	攻击方向	17
4.5.5	攻击方式	17
4.5.6	攻击所用资源	17
4.5.7	元特征扩展	18
5	扩展钻石模型	19
5.1	社会政治	20
5.1.1	持久对手关系	20
5.1.2	网络受害者学	23
5.1.3	共享威胁空间	24
5.2	技术	24
6	上下文指标	25
7	支点分析	26
7.1	‘中心’方法	26
7.1.1	以受害者为中心方法	26
7.1.2	以能力为中心方法	28
7.1.3	以基础设施为中心方法	29
7.1.4	以对手为中心方法	29
7.1.5	以社会政治为中心方法	29
7.1.6	以技术能力为中心方法	30
8	活动线	30
8.1	对手过程	36
8.2	假设分析支持	36
8.3	行为攻击图	39
9	活动组	40
9.1	步骤1：分析问题	42
9.2	步骤2：特征选择	43
9.3	步骤3：创建	46

9.3.1 活动组创建示例	47
9.4 步骤4: 增长	47
9.5 步骤5: 分析	50
9.6 步骤6: 重定义	50
9.7 活动组族	50
10 规划和博弈	52
11 未来工作	55
12 结论	56
参考文献	57

图示列表

1 钻石模型事件 9

2 扩展钻石模型事件 19

3 对手-受害者关系 21

4 持久度谱 22

5 使用钻石模型支点分析示例 27

6 钻石模型活动线示例 31

7 钻石模型对手过程示例 37

8 活动攻击图示例 39

9 活动组创建 47

10 活动组增长 48

11 钻石模型/杀伤链行动矩阵 53

1 引言

入侵分析学科自发现第一次入侵以来就一直存在¹，外部黑客和内部不法人员大多情况下都是狡猾隐秘地进行渗透和攻击，而入侵分析人员和系统管理员则致力于发现、理解和阻止他们的恶意行为。自该学科诞生以来，问题几乎没有改变过：是谁，干了什么，在何时，在何地，为什么以及结果如何。从以往经验来看，这些问题为事件响应提供了信息以解决基本的问题，但防御者缺乏必要的活动记录、综合推理和关联性分析模型和框架，以回答一个日益重要的问题：作为协同活动的一部分，对手还会回来吗？然而，最终这个问题导致组织背离了战术缓解而倾向战略缓解，因此增加了缓解策略的有效性和与对手进行作战的成本。

本文根据分析人员多年的工作经验提出了一种入侵分析新模型，该模型抛出了一个简单的问题：“我们工作的基本方法是什么？”，它得名于钻石模型，是因为它将恶意活动的最基本方面简单地组织成钻石的形状。我们的模型首次建立了一种将科学原理应用于入侵分析的形式方法：在可测量、可校验和可重复方面提出了一种简单、常规和全面的活动记录、组合和关联分析方法。这种科学的方法及其简易性提高了分析的有效性、效率和准确性。

我们的模型既简单又复杂，既抽象又具体，可用于分析内部和外部威胁。抽象地说，分析人员能很容易理解这个模型，使其在日常的工作中发挥作用。该模型是本体论的基础²，并提出了一个框架，在此基础之上发现新活动、最大化中枢分析时机、关联和合成新信息，并随着时间的推移追寻对手，同时健全沟通和信息记录。

具体来说，该模型是一个数学框架，应用了博弈论、图论和分类/聚类理论来改进分析和决策。它有几个好处：可校验的分析假设能确保分析结果的可重复性和准确性，生成假设更容易，跨事件自动关联，可以快速可信将事件分类到对手活动中，以及可以在规划和实施缓解策略时预测对手活动。最终，这种形式将导致该模型能够将相关情报集成到网络防御能力中，从而轻松地发现新的对手基础设施、能力和攻击过程。

最重要的是，该模型是目标明确的通用模型，灵活可扩展。它准确地定义了入侵分析和

¹ 在本文中，“入侵”一词是指针对计算机系统和网络的所有恶意和非法活动。

² 该模型不提供新的本体论、分类法、共享格式或协议，但根据其基本性质，应该构成这些模型的基础，[3]中的其他人也支持此观点。

对手行动的基本概念，这些属性增强了模型的实用性，使其能够成长并包含新的思想和概念。

2 相关工作

在入侵分析中，我们与分析人员和专家站在一起，如Stoll[4]、Bellovin[5]和Cheswick[6]，他们发现并记录恶意事件，但是他们几乎没有正式的培训 and 工具。他们通常依靠大量的数据打印输出来分析活动，并且仅仅靠直觉和超高的技术能力。他们最初的技术文章和光辉事迹，让许多分析人员走上了猎捕对手的道路，现代入侵分析人员通过诸如Honeynet项目[7]等杰出和创新的工作延续了这一传统。

Northcutt[8]和其他安全研究人员通过展示具体的威胁活动示例，并为学生提供了解对手的工具和情报的机会，以增强分析训练。来自SANS[9]等组织的实操培训现在是分析技术传播的重要来源。

尽管这些报道、论文、书籍和课程为教授入侵分析技术提供了坚实的案例，但它们没有提供巩固该学习过程所必需的科学方法。如果没有基础模型（正式的或非正式的）来解释分析人员如何评估和理解恶意活动，那么入侵分析技术就很难取得进步。

其他工作偏离了入侵分析和情报驱动的网络防御的使用。Amann等人在[10]中如实地指出，“在没有外部背景的情况下，可靠地报告如今的复杂攻击情况变得越来越困难。然而，不幸的是，如今的入侵检测系统（Intrusion Detection Systems, IDS）无法轻易地集成情报.....”，他们的工作显着提高了入侵检测系统实时集成外部环境和威胁情报的能力，从而提高了检测成功率。这是未来缓解攻击的关键能力，钻石模型通过确定分析人员如何有效、高效和准确地结合外部环境和情报来补足了这一能力，从而提高了检测能力。

“杀伤链”提供了一个高效重要的对手行动模型，它直接影响到解决措施的决策[11]。我们的模型集成了“杀伤链”模型的分阶段方法，并通过扩展视角来补充杀伤链分析，这种视角提供了所需的分析粒度和入侵活动之间复杂关系的表达。这样就可以表示全部认知范围，而不只是活动中可观察的指标。此外，我们的模型提供了一种形式化的数学方法来进行有效的图形分析和分组（例如聚类/分类），以解决许多类别的分析问题。该特性允许该模型支持许多互补的战略规划框架，如联合作战情报环境准备（JIPOE）[12]，行动过程基础[11]，主动防御算法和模型（ADAM）[13]，并且可能使用诸如[14]之类的先进计算技术进行更“前沿”的

策略规划。

传统的攻击图试图为一组给定的受保护资源生成所有可能的攻击路径和漏洞，以确定最具成本效益的防御方式和最大程度的保护措施。攻击图源于Schneier的“攻击树”，逐渐发展成为一个实用的脆弱性分析工具，可以帮助制定有效的纵深防御策略[15]。直到2005年，攻击图在可伸缩性、度量性和可用性方面都面临重大问题[16]。然而，实际规模的网络的可扩展性[17,18]、度量性[19]和可用性方面[20]已经取得了进步。我们的模型定义了一个新的情报驱动的攻击图，称为活动线，同时将情报和传统攻击图组合成活动攻击图。活动攻击图将传统的脆弱性分析与对手活动的知识相结合。它们将已发生的情况与潜在的和首选的攻击向量相结合，从而实现更有效地攻击分析和策略制定。这最终可以更有效地分配防御资源。此外，在先前工作[21]已经直接在入侵检测系统中显示了攻击图的适用性。，这使得在我们的模型中提出的活动线和对手进程可以直接在入侵检测系统中得到实现。

许多系统，语言和分类法可以让分析人员记录恶意活动并共享特征，例如像这样的特征组[22,23,24,25,26,27,28,29,30,31,32]。我们的模型没有涉及到本体论、分类学或者共享协议，然而在最近对网络安全本体的一项调查中，我们的模型作为调查基础被引用，并建议依据我们的模型来整合现有的网络安全本体并构建未来网络安全本体[3]。此外，我们的模型支持这样一种观点，即要真正整合网络威胁情报，我们必须避免将复杂而深入的关系活动表示为简单的技术指标列表。我们认为，为了实现战略性的缓解，对于入侵活动，必须在保留关键复杂关系的同时记录和共享整合的非技术性内容。

3 钻石模型概述

该模型最简单的形式（图1）描述了攻击者在某些基础架构上针对受害者的功能部署。这些活动称为事件，是此模型中的原子特性。分析人员或机器随着事件的被发现和调查进展来填充模型的顶点。顶点之间通过表示特性之间的自然关系的边进行连接，从连接边和顶点内不同的角度，分析人员可以了解更多关于对手行动的信息，并发现新的能力、基础设施和受害者。

一个事件只是持续攻击活动中的一个步骤，对手必须执行该步骤才能实现其目标。因此，事件是由对手-受害者成对进入活动线的进行阶段顺序的，其中活动线表示对手行动的流程。

事件和活动线都是深入理解恶意活动的必要元素，因为更有效的缓解策略“需要对入侵本身进行新的理解，事件不要只看成单一事件，而意识到它是阶段性过程的一部分。” [11]

一旦建立了活动线，就可以跨活动线关联事件，以识别对手的活动，并合并为活动组，以识别具有共同特征的类似事件和威胁。这些活动组可用于事件的自动关联，也可用于博弈和规划缓解策略以及建立对抗对手的战略上的缓解计划的场景。

上述术语和概念将在下面的章节中进一步描述和讨论，下面从模型的原子元素——钻石事件开始。

4 钻石事件

公理 1 对于每一个入侵事件，都有一个对手通过在基础设施上对受害者使用能力来产生结果，从而朝着预期目标迈进。

一个事件定义为限制在特定阶段的离散时间范围的行动，在这个阶段，需要外部资源的对手在某些基础设施上使用一种能力和方法来对付具有给定结果的受害者。当然，不是所有的特征都需要知道才能创建一个事件。在许多情况下，大多数特征都是未知的，只有随着新的事实被发现以及收集了更多的数据才能了解清楚所有特征。

核心特征 一个事件的核心特征有：对手、能力、基础设施、受害者。

元特征 元特征包含：攻击时间戳（起始时间和结束时间），攻击阶段，攻击结果，攻击方向，攻击手段，攻击所用资源。元特征用于在活动线（见8节）内对事件进行排序，以各种方式对事件进行分组，并在可能的情况下捕获关键信息。

置信度 每个事件特征，无论是核心特征还是元特征，都有一个相关的置信值。这个值是有意保留的，因为每个模型的实现对于置信度都有着的不同理解。此外，置信度可能是有多个值的函数，比如分析结论的置信度或数据源的准确性。如有必要，置信值也可以作为子元组逐项列出，以便更好地捕获置信度的各个独立元素。

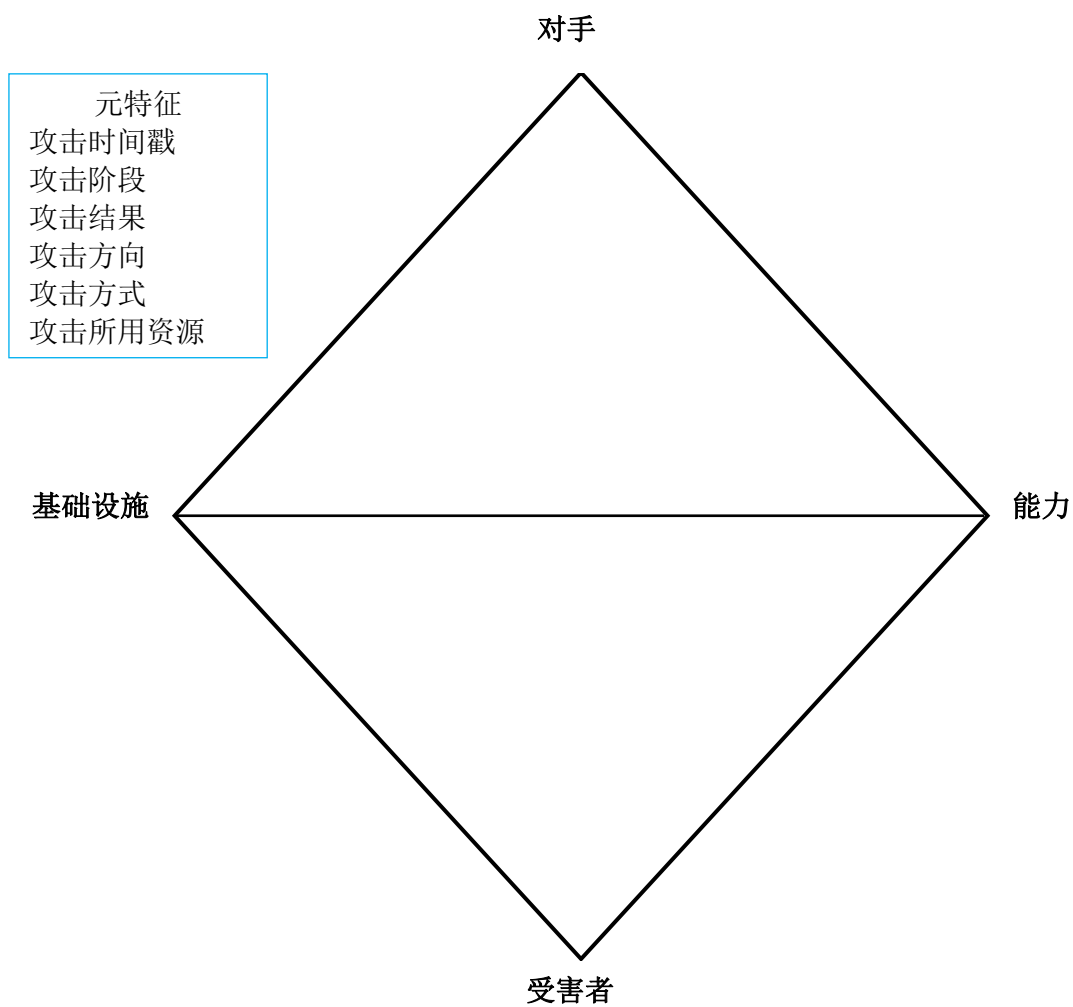


图1：入侵分析钻石模型，包括入侵事件的核心特征：对手、能力、基础设施和受害者。核心特征通过边缘进行链接，以表示特征之间的基本关系，这些特征可以通过分析来进一步发现和获取恶意活动的信息。元特征也列出了，虽然不是核心特征，但强调了它们在高阶分析、分组和规划功能中的重要性。

该模型的一个好处是，它提供了在每个事件中应该出现的有效（但不一定全面）的特征列表。因此，在用所有可用信息记录一个事件之后，现在任何空的特征都被识别成信息缺口，该模型鼓励通过额外的支点来弥补这些缺口。

事件 E ,定义为带有标签的 n 元组，其中元组中的每个元素都有一种特征的信息，其结合了一个互相独立的置信值³

$$E = \langle \langle \text{Adversary}, \text{Confidence}_{\text{adversary}} \rangle, \langle \text{Capability}, \text{Confidence}_{\text{capability}} \rangle, \langle \text{Infrastructure}, \text{Confidence}_{\text{infrastructure}} \rangle, \langle \text{Victim}, \text{Confidence}_{\text{victim}} \rangle, \langle \text{Timestamp}_{\text{start}}, \text{Confidence}_{\text{timestamp}_{\text{start}}} \rangle, \langle \text{Timestamp}_{\text{end}}, \text{Confidence}_{\text{timestamp}_{\text{end}}} \rangle, \langle \text{Phase}, \text{Confidence}_{\text{phase}} \rangle, \langle \text{Result}, \text{Confidence}_{\text{result}} \rangle, \langle \text{Direction}, \text{Confidence}_{\text{direction}} \rangle, \langle \text{Methodology}, \text{Confidence}_{\text{methodology}} \rangle, \langle \text{Resources}, \text{Confidence}_{\text{resources}} \rangle \rangle$$

为了增加灵活性，可以将基本元组扩展为嵌套有序对的层次结构（为简单起见，此处称为子元组），以进一步定义特定特征并捕获知识以备将来的关联。

这是一个扩展受害者特征的示例，例如：受害组织，主机的IP地址，主机名，被利用的应用程序以及该应用程序的TCP端口：⁴

³ 事件是一个大小可变的 n 元组，而不是一个固定大小，因为模型不仅限于这里定义的特征，而且可以扩展到包括其他感兴趣的元素，如扩展钻石模型5中的元素。一旦一个组织定义了它们的完整特征集，就将为该特定实例正式定义元组大小。

⁴ 注意，在这个例子中，受害者的每个子特征都有一个独立的置信值，但是也可以实现一个模型，其中只用一个置信值简单地应用于所有子特征。

$$\begin{aligned} \langle \text{Victim}, \text{Confidence}_{\text{victim}} \rangle = & \\ & \langle \langle \text{Organization}, \text{Confidence}_{\text{organization}} \rangle, \\ & \langle \text{HostIP Address}, \text{Confidence}_{\text{IP}} \rangle, \\ & \langle \text{Hostname}, \text{Confidence}_{\text{Hostname}} \rangle, \\ & \langle \text{Application}, \text{Confidence}_{\text{Application}} \rangle, \\ & \langle \text{TCP Port}, \text{Confidence}_{\text{TCP Port}} \rangle \end{aligned}$$

出于分析目的，事件也可以理解并表示为图1所示的图。在这种形式中，边缘代表事件特征之间的自然关系，并通过支点（在第7章中进一步描述）从该特征的角度识别通常可见或可发现的内容。核心特征（对手、能力、基础设施和受害者）构成一个无向的简单顶点图。因此定义了一个图形化的事件E：

$$\begin{aligned} E_{\text{vertices}} = \{ & \text{Adversary}, \\ & \text{Infrastructure}, \\ & \text{Capability}, \\ & \text{Victim} \} \end{aligned}$$

$$\begin{aligned} E_{\text{edges}} = \{ & \{ \text{Adversary}, \text{Capability} \}, \\ & \{ \text{Adversary}, \text{Infrastructure} \}, \\ & \{ \text{Infrastructure}, \text{Capability} \}, \\ & \{ \text{Infrastructure}, \text{Victim} \}, \\ & \{ \text{Capability}, \text{Victim} \} \} \end{aligned}$$

4.1 对手

公理 2 存在一系列的对手（内部人员、外部人员、个人、团体和组织），他们试图破坏计算机系统或网络以进一步扩展其意图并满足他们的需求。

对手是指利用针对受害者的能力来实现其意图的人或组织。对手的能力水平通常是难以捉摸的，对于大多数事件来说，这个特征可能是未知的——至少在发现的时候是未知的。

在分析事件技术方面的大多数时候，我们只是简单地将对手操作者称为对手。然而，通过帮助构建对手和受害者之间的关系，区分对手操作者和对手消费者对于理解意图、归属、灵活性和持久性是很重要的。因此，我们发现这些区别尤为重要⁵：

对手操作者 这是实际进行入侵活动的“黑客”或人员。

对手消费者 这个实体受益于在入侵中进行的活动。它可以与对手操作者相同，也可以是单独的人或组。

举个例子，一个资源充足的对手消费者可以在不同时间或同时指挥不同的操作者，每个对手操作者都有自己的能力和基础设施，以实现共同或单独的目标⁶。而相比之下，单一的对手操作者拥有的能力和基础设施就相形见绌，同时单一的对手操作者也缺乏绕过简单防御措施的能力。

如果对手操作者及其客户作为一个独立实体存在，了解其动机和资源配置将有助于衡量其对受害者的真正威胁和风险，从而更有效地规避风险。了解这些动机是社会政治需要，稍后在扩展的钻石模型（见第5节）中解释。

4.2 能力

能力特征描述了在事件中对手使用的工具和/或技术。模型的灵活性使得其能够准确地描述能力。我们希望能力被广泛理解并包含所有影响受害者的手段，从手动的“简单”方法（如手动密码猜测）到最复杂的自动化技术。

能力水平 无论受害者如何，个人可以利用的所有漏洞和隐患都被视为其能力。

对手武器库 一个对手的全部能力，也即他们个人能力的综合能力，是对手的武器库。

⁵ 虽然在整个工作中建议了一些特性区别和类别，以便对日常分析有用，但是没有对完整性的声明，也没有建议这些区别构成一个本体或分类法，因此模型也不需要这些区别。

⁶ 由更高权限组织的各种活动可以作为活动组系列在模型中进行建模和组织（见9.7节）。

如果能力水平已知，则应将其记录为能力的子元组，以及活动攻击图（见8.3节）上的潜在路径。通过活动组分析（见9.5节），对手能力的初始记录可以随着时间的推移而增多，最终达到对其武器库的了解。这些宝贵信息在缓解攻击和防御规划中具有重要的作用，使得防御一方能够预测对手的行动和反应。

4.2.1 命令与控制

命令与控制（C2）是控制人员对资产进行权限管理的操作。在入侵分析中，这意味着来自对手的通道，通信结构，信号，协议和内容（例如，获得访问权、故意删除访问权、窃取数据、发送攻击包）都能够影响对手实现目标的步伐。

虽然命令与控制可以采取多种方式，但最终都取决于使用的能力，在支点分析方面，分析人员以C2为中心，发现基础设施和受害者之间的通信。因此，就我们的模型而言，命令与控制最好理解为能力的一个子特征。

4.3 基础设施

基础设施特征描述了对手用来提供能力、维持控制（例如，命令与控制/C2）以及从受害者那里得到结果（例如，窃取数据）的物理和/或逻辑通信结构。与其他特征一样，基础设施可以是狭义的，也可以是广义的。示例包括：互联网协议（IP）地址、域名、电子邮件地址、甚至是街道对面的手机语音信号灯闪烁的莫尔斯电码、停车场中的USB设备和插入工作站中的USB设备，或附近的监听站收集的来自硬件（如Van Eck Phreaking）破解信号。以下基础设施角色分类对于大多数入侵分析目的来说是合理的。

类型一基础设施 完全由对手控制或拥有的基础设施，或他们在物理上可以接近的基础设施。

类型二基础设施 由（知情或不知情）的中间人控制的基础设施，通常，这些受害者被视为对手的基础设施，它有助于混淆对手的来源和属性。这类基础设施包括僵尸主机，恶意程序感染的服务器，恶意域名，跳板机，受感染的电子邮件账户等。

服务提供商 为对手提供一二类基础设施可用的关键服务的组织（如互联网服务提供商、域名注册商、电子邮件提供商）

4.4 受害者

受害者是对手借助其能力达到漏洞利用的目标，与其他特征一样，可以以任何必要和适当的方式描述被害人：组织、个人、目标的电子邮件地址、IP地址、域等。单独定义被害者用户模型及其资产非常有用，因为它们提供了不同的分析功能。被害者角色在非技术分析中很有用，如网络受害者学（见5.1.2节）和社会政治中心方法（见5.1节），而被害者资产与常见的技术方法有关，如脆弱性分析。

受害者身份 受害者身份是指资产被利用和攻击的个人和组织。包括组织名称、人名、行业、职务、兴趣等。

受害者资产 受害者资产即攻击面，包括网络、系统、主机、电子邮件地址、IP地址、社交网络帐户等，对手根据攻击面来施展其能力。受害者资产通常存在角色控制和内外可见性的情况，但仍可能被对手盯上。常见的例子包括网络邮件帐户和基于云存储的数据。

在一个事件中，受害者资产可以是最终目标（例如受害者），然后在以后的事件中用作基础设施（如前文4.3所述，可能是第二类基础设施）。在这情况下，分析人员必须始终明白行动的首次目标未必是受害者。

4.4.1 系统漏洞

公理 3 每个系统，以及每个受害者资产，都有漏洞和风险。

对手利用公理3定义的漏洞和风险来实现其意图。钻石模型的灵活性将这些定义为受害者的子特征。这些可以概括地描述为“缺乏对用户的教育，致使用户点击电子邮件中的超链接”，或者具体描述为符合记录要求的CVE[35]。

受害者脆弱性 易受到攻击的受害者的漏洞集合被称为受害者的脆弱性

在我们的模型中，受害者脆弱性列表能很容易表示为受害者的子元组。与能力水平和对手武器库（见第4.2节）相比，这一信息非常有价值，可用于确定缓解方案。与能力水平一样，还可以用活动攻击图（见第8.3节）交替或联合描述。

4.5 事件元特征

事件元特征稍微扩展了模型，包括了一些非关键但重要的钻石事件元素。下文所述的一些元特征是我们觉得比较有用的，但是我们的模型并不局限于这些元特征。那些实现或扩展我们模型的人可能希望能添加额外的元特征来获取与事件相关的其他关键信息元素。

4.5.1 攻击时间戳

每个事件都有发生的日期和/或时间。它可以根据需要具体化，也可以表示事件开始和停止的范围。时间戳是归类恶意行为的一个组成部分，因为随着时间的推移，对手发生变化的可能性就会增加，那么就会降低对于信息的置信度（也就是衰减函数），此外，时间戳与随时间变化的对手事件集合结合，可以产生其他独特的分析形式，如[6]中所述的建立的周期性和对于对手生活模式的推断。

4.5.2 攻击阶段

公理 4 任意恶意行动都包含两个或多个阶段，必须连续成功地执行这些阶段才能实现既定目标。

恶意行为不是在单个事件中发生的，而是在两个或多个事件中发生的。其他一些安全研究员提供了充分的证据支持所有入侵行为应被视为事件链的观点[11,36,15]。举个例子，首先，对手必须找到受害者（通常使用研究和/或扫描），然后发现易受攻击的主机，接着是漏洞利用、建立命令控制连接，最后是某种行为。通常，在入侵的各个阶段使用的能力不同，有时使用的基础设施也不同。不过至少，首先必须确定受害者身份（可以简单地选择一个随机IP地

址），然后再执行之后的行动。

我们的模型可以适用对手行动的任何阶段（如[11,36,15]）⁷，这是一个比较重要的特性，因为事件所处的阶段表明了它在行动线中的位置（见第8节）。

尽管公理4表明每个行动都包含一系列阶段，但是还没有共识或证据表明存在一组满足所有恶意行为特征的阶段。事实上，许多的多阶段活动定义的存在表明了有不同的情况，例如[36, 11]中给出的示例。因此，我们假设钻石模型的使用者可以为某些活动定义一些非必要的阶段。

在形式上，我们定义 P 为有序 n 元组的阶段集，其中 n 是使用该模型的用户定义为必要且足以描述所有可能事件的阶段数，每个事件由唯一的有序阶段集 P 表示。

其中：

$$P = \langle p_1, p_2 \dots p_n \rangle$$

- $n \geq 2$ （由公理4可知至少存在两个及以上的阶段）。
- p 为对手行动中的某个阶段。
- p_1 是对手行动中的第一个阶段。
- p_{n+1} 是执行阶段 p_n 之后的阶段。

4.5.3 攻击结果

虽然对手行动的结果和后置条件并不总是已知的，或者在已知时具有很高的置信度了，但捕获它们是有用的。纵观对手的作战行动对于确定攻击者具有特定能力的成功率或针对受害者们的子集的成功率有很大的帮助。后置条件的集合能够提供对于对手意图更宽广的观点。有几种能够记录结果的方法。一种方法是使用三元组：（成功，失败，未知）。另外一种方法是通过信息安全基本属性将其分类：机密性受损、完整性受损和可用性受损。还有另外一种方法是可以通过记录事件导致的所有后置条件，例如（在侦察阶段）获得的目标信息或随

⁷ 确定每个事件的阶段对于维护知识和关联事件并不重要，但对于缓解计划和杀伤链分析非常有用。

后在伪装攻击中有用的泄露的密码。此外，还可以使用现有的攻击结果分类法，如[26]中Cohen描述的类别。

4.5.4 攻击方向

当考虑缓解方案和检测位置时，事件的方向性非常重要。这个元特征在描述基于网络的事件时通常很有用，但也可以用于描述基于主机的事件。这个特征大概有七个潜在选项：受害者到基础设施、基础设施到受害者、基础设施到基础设施、对手到基础设施、基础设施到对手、双向、或未知方向。通过维护这些信息并考虑对手随着时间变化的行动方向，更好地决定使用仅外部、外部或内部的检测和缓解措施的某种组合最适合对抗对手。

4.5.5 攻击方式

攻击方式元特征使得分析人员能够描述一般的行动类别，例如：鱼叉式钓鱼邮件、内容分发攻击、syn泛洪攻击、端口扫描等。与其他特征一样，攻击方式可以根据需要进行多个定义。比如说，带有恶意程序的恶意鱼叉式钓鱼邮件可以归为“鱼叉式钓鱼邮件攻击”或者“内容分发攻击”。而带有超链接的鱼叉式钓鱼邮件可以归为“鱼叉式钓鱼邮件攻击”或者“用户重定向攻击”。这种方法可以更好地对事件进行分类，并可以针对单个对手和多个对手进行独立于指标的事件比较，以达到分组和缓解的目的。

可以很容易地将几个现有的分类法合并到这个特征中，从而减少工作量，并提高与现有框架的互操作性，比如Snort入侵检测系统的classtype规则[37]和许多更正式的研究[25、29、26、28]。

4.5.6 攻击所用资源

公理 5 每个入侵事件在成功之前都需要一个或多个外部资源。

资源元特征列出了攻击事件需要的一个或多个外部资源。资源应被广义理解为事件以及每个核心特征和元特征所依赖的任何和所有辅助元素。当考虑到资源约束和重心缓解策略，以及知识鸿沟和假设检验（如下文第8.2节所述）的识别时，这个元特征变得很重要。

显然，这个元特征可能包含许多元素数量。然而，和其他特性一样，钻石模型不需要完整性，具备充分性就可以了。因此，一个组织只需列举对其特定用途有效的所需资源。

示例资源如下：

- 软件（例如：metasploit、操作系统、虚拟化软件）
- 知识（例如：如何使用metasploit,以及在何处获取漏洞信息）
- 信息（例如：用于伪装的用户名以及密码）
- 硬件（例如：工作站、服务器、调制解调器）
- 资金（例如：购买域名的资金）
- 设施（例如：电力，避难所）
- 访问权限（例如：从源主机到受害者（反之亦然）的网络路径、可路由的IP地址和来自Internet服务提供商（ISP）提供的网络访问）

4.5.7 元特征扩展

上面已经描述了几个可以很好的集成到模型中的元特征，当然恶意入侵事件还有很多其他的元特征，可以根据组织的需要考虑包含这些元特征：数据源（捕获或检测到事件的数据源）、作者（事件分析人员/作者）、检测方法（检测恶意事件使用的工具，技术或能力）、检测签名（检测恶意事件的签名或启发式签名）等。添加额外的元特征能够增强模型，使得用户、分析人员和组织维护与事件相关的重要信息以供将来使用（例如更有效的来源，鼓励发现和原创，优化分析，理解置信区间，质量控制等）

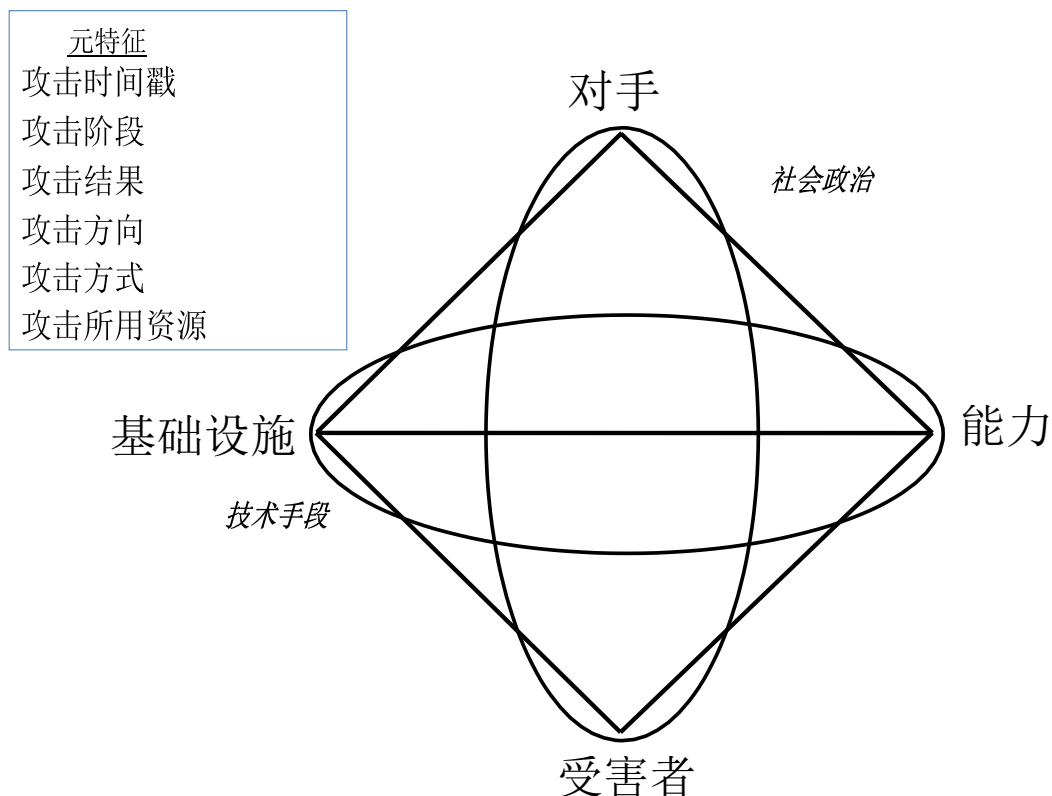


图2：扩展钻石模型包含了独特的社会政治和技术手段特征，这些特征突出了对手-受害者（通过对手的社会政治需求、愿望和动机以及受害者满足这些需求的能力）与能力-基础设施（通过用于实现其通信的技术）之间的特殊关系。

5 扩展钻石模型

如前所述，钻石模型很容易通过包含其他必要的特征来进行扩展。如图2所示，任何入侵活动的另外两个基本特征是：决定对手-受害者关系的社会政治元特征，以及实现基础设施和能力的技术能力元特征。这两个独特的特征通过叠加另外两个无法分割的特征定义了如下关系：一个跨越了对手-受害者轴，另一个跨越了能力-基础设施轴。

5.1 社会政治

公理 6 无论接触距离远近和时间长短以及是否直接或间接，对手和受害者之间始终存在着关系。

对手-受害者关系是以生产者-消费者关系为基础的，这一关系是由对手的社会政治需要和愿望（例如，创造收入、在黑客社区获得认可，成为公认黑客，增加商业利润）⁸，这种关系表示对手的需要以及受害者满足对手意图（例如商业间谍、传统间谍、民事欺诈、拒绝服务攻击、网站篡改）。受害者无意中提供了一种“产品”（例如，在僵尸网络中计算资源和带宽，作为宣传和工业的目标或者用于商业间谍的商业敏感信息，用于欺诈的财务信息和用户名/密码），而对手“消费”了他们的产品。

意图 虽然意图对于理解入侵活动很重要，并且应该作用于防御决策，但并不是钻石模型的基本顶级元特征，而是更适合作为社会政治子结构中的一个特征，这进一步可以假设更高的需求和愿望。

5.1.1 持久对手关系

结合公理2和公理6可以推出：存在对手莫名其妙就和他的受害者建立了关系的情况，然而并不是所有对手-受害者关系都是一样的。一些对手在进行“入侵掠夺”时只关心当前可用的访问权限和数据，并不担心随时可能失去访问权限。不过也有些对手即使面对强大的保护措施仍坚持不懈对受害者进行攻击，也有一些对手甚至对那些抑制他们行动的人进行报复。对手坚持从受害者处获取访问权限和/或信息是一种很重要的对手-受害者关系的描述。

因此，考虑到持久和非持久关系，可以假定以下公理：

⁸ “社会政治”一词被谨慎地选作对大多数对手的广泛需求和愿望的分类，包括但不限于个人、伙伴关系、松散组织的集体、等级集团、非国家和国家行动者。他们的需要可以从社会和政治两个方面进行广泛描述。有人可能会说，政治是社会需求的延伸，是个人在权威下组织下以满足集体欲望的延伸。然而，我们认为，将这两个术语分开能够在个人/非国家群体需求与权威（如政府或军队）需求之间取得最佳的对话平衡，以及这些需求如何用于受害者的选择，从而做出缓解决定。

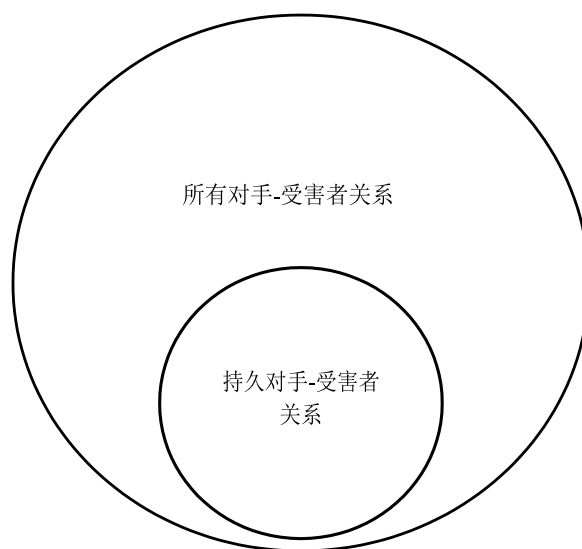


图3：上面的维恩图说明了由公理2和6定义的所有对手-受害者关系的集合，以及由公理7定义的持久对手关系的子集合。

公理 7 存在一组对手的子集，这些对手有动机、资源和能力在抵抗缓解抑制的同时，对一个或多个受害者维持相当长时间的恶意影响。此子集中的对手-受害者关系称为持久对手关系。

持久对手 持久对手是在特定的对手-受害者关系中满足公理7的对手。图3说明了由公理2和公理6定义的所有对手-受害者关系与由公理7定义的持久对手子集合之间的关系。将对手-受害者关系置于其中一个集合中取决于对公理7的符合度。

没有必要因为对手对一个受害者坚持不懈，所以认为他们对所有受害者都坚持不懈。例如，在某次行动中，对手获得访问权后确定受害者没有价值，并且不考虑持久性控制而离开。然而，在另一个活动中，对手可能会坚持更长的时间以获得更多的价值。从另一个角度来看，受害者可能是多个对手的宿主，其中一些对手可能是持久的，而另一些则是非持久的。因此，持久性或非持久性是由特定的对手-受害者关系对决定的。

Fleeting Enduring

图4：持久度谱表明，并非所有对手的持久关系都是相同的，而是落在短暂和持久之间的谱上。当一个特定的对手-受害者关系落在频谱上时，它是许多因素的作用的结果，并且随着时间的推移而变化。

此外，持久性既不是二值的，也不是静态的。众所周知，许多持续的入侵可以通过技术措施来实现防御，如Stoll[4]，Cheswick在“Berferd”[6]中阐述的一些对手抵制技术措施，甚至包括尝试公开羞辱对手的方法。在“Berferd”的案例中，通过给黑客的母亲打电话，最终实现了缓解。因此，根据持久性程度不同，我们提出以下推论：

推论 1 根据对手-受害者关系的基本原理，存在不同持久性程度的对手。

持久度表明了对手的动机和技术能力，以及对手为维持其效果而付出的努力和资源。持久度表现在短暂到耐久之间的光谱上，如图4所示，在许多情况下，持久度决定了防御者抵抗攻击所需的努力和资源的数量。对手的动机和能力越强，他们对缓解措施的抵抗力越强，这就意味着一种更持久的关系正在向频谱的右侧移动。

传统上，缓解措施仅限于围绕对手能力而采取的技术手段，对其动机和资源的影响很小，导致对手在从受害者身上撤离后很快返回。通过使社会政治关系及其相关需求和愿望成为恶意活动的关键部分，钻石模型使非传统领域（如心理学、犯罪学、受害者学、营销、消费者行为和经济学）的应用得以扩大缓解的多样性。特别是，它支持对手的决策以及他们的感知偏好，这些都可以被控制和影响，有利于防守方除了传统的技术之外的选择。

以下是决定持久性程度的对手-受害者关系的一些要素：

- 与其他需求相比，受害者满足对手需求的相对强度。
- 对手认为持续攻击的风险。
- 对手维持攻击的成本。

- 受害者满足特定需求的独特性。
- 受害者对需求的持续满足。
- 防守者为抵抗攻击者的持续性而付出的努力和资源。

对于持久或非持久的对手关系，持久度频谱上的位置对于每个对手-受害者对都是唯一的。为了便于参考和分析，在不忽略频谱所代表的复杂性和连续性的情况下，我们通常考虑频谱上的两类受害者：机会受害者和利益受害者。

机会受害者 这类受害者是对手行动中的一种消耗性商品，在这种行动中，攻击者很可能不会注意到失去访问权，也不会导致对手花费资源重新获得访问权。这一类的受害者落在持久性光谱的左侧，朝“短暂”以及非持久性关系方向靠拢。这些受害者很可能是最初的目标，因为他们很脆弱，随时都可以被利用。

利益受害者 这类受害者是一种非消耗性商品，在这种商品中，持续可访问为对手提供了足够的价值，失去访问权会引起对手注意，并且对手将花费资源重新获得对该或相关受害者的访问权。这类受害者落在持久度谱的右边，朝向“持久”。

重要的是，持久的对手-受害者关系并不是一成不变的——而是可以改变的。仅仅因为一个受害者开始时是短暂的，而一个机会受害者并不意味着他们以后不能改变。例如，如果一个受害者最初被一种自我传播的蠕虫所利用，但对手发现受害者的价值高于消耗性商品，那么他们可能会成为一个利益受害者，沿着持久度谱走向“持久”。

5.1.2 网络受害者学

我们的模型独一无二的把受害者和对手放在一个对等的空间里，突出了两人之间通常无法言喻的关系。此外，随着我们的模型通过活动线（8）和活动组（9）扩展到包括许多对手和受害者，我们可以开始从犯罪学和受害者学汲取专业知识，从而提出重要问题，如：

- 为什么某个特定个体受害？
- 是否有一组共同的受害者？

- 受害者有共同特点吗？
- 我们能从受害者集合中推断出意图吗？
- 谁可能是其他但未知的受害者？
- 谁有需要和意图伤害这些组织？

重要的是，通过一个更好的受害者模型，我们可以开始研究通过降低受害者的吸引力和预测未来的受害者来对抗对手的方法。这使得一个组织能够适当地最大化检测资源，就像一个侦探关注的是高风险人群和集中犯罪区域，而不是巡逻随机区域一样⁹。

最近的水坑攻击¹⁰说明了对手如何利用这个概念来描述他们的受害者，以便在最赚钱的地方进行攻击。例如，在2013年4月，一个在藏族激进分子相关网站上试图利用易受攻击浏览器的访问者[40]。但是如果有效结合社会政治特征与以受害者为中心的方法（见7.1.1节），则可以预测一些水坑攻击的对象，并进行有针对性的检测/缓解，先发制人地阻止恶意活动。

5.1.3 共享威胁空间

如果两个或两个以上的受害者具有足够的相同特征，能够满足一个或多个对手的需求，那么他们就处于“共享威胁空间”。早期识别共享威胁空间是战略和主动缓解攻击的基石。例如，针对一个成员的目标攻击使集体威胁空间能够预测和预测未来的攻击。此外，与最有可能受到类似对手威胁的人分享威胁情报更为有利。

5.2 技术

除了社会政治元特征外，技术元特征还突出了一种特殊的关系，跨越了两个核心特征：能力和基础设施。这代表了连接和启用基础设施的技术以及操作和通信的能力。

⁹ 一项有趣的犯罪学研究表明，城市社区内树木覆盖率的增加在统计学上与犯罪率的减少相关[38]。是否可能存在入侵活动的潜在并行相关的因素？

¹⁰ 水坑攻击是一种方法论，在这种方法论中，对手会破坏合法的网站，他们认为他们的目标群体将访问这些网站，从而利用这些网站。这个比喻和术语是从狮子埋伏在一个水坑里埋伏猎物中提取出来的[39]。

例如，如果安装的恶意程序解析域并通过HTTP进行通信，则使用的技术有：Internet协议（IP）、传输控制协议（TCP）、超文本传输协议（HTTP）和域名系统（DNS）。通过分析技术及其潜在的异常/误用，分析人员发现了新的恶意活动，而不管潜在的基础设施和能力如何（也称为以技术为中心的方法7.1.6）。此外，了解对手行动中涉及的技术有助于确定最合适的探测位置、数据类型和能力。

6 上下文指标

指标是系统和分析人员用来检测对手行动的信息元素。在正常的业务过程中，指标被加载到检测系统中，从而警告分析者潜在的对手行动。传统指标仅限于技术细节。一些人将这些扩展到包含额外的元数据[32]。然而，现在是时候将指标扩展到包括非技术性、行为性和概念性的元素，这些元素可以增强检测能力但不能直接由自动检测实现。

上下文指标 上下文指标是一种信息元素，它被放置在对手作战的上下文中，以丰富探测和分析。钻石模型衍生的上下文指标确保元素之间的关系及其作用被保留，并且分析概念如对手需要和意图被完全结合，从而产生更完整的上下文。

例如，在传统的指示方法中，对手的基础设施IP地址是一个常见元素。使用我们的模型作为本体论的基础，这个IP地址可以放在上下文中，不仅为分析人员提供对对手基础设施（可能是检测警报）的了解，而且还提供对先前被破坏的受害者类型/类别以及对手试图破坏的信息项（例如商业计划文件、知识产权）。利用这一增强的知识，分析者现在不仅可以检测和确认入侵（可通过传统指标获得），还可以确定他们是否是对手行动的一部分、对手可能瞄准的信息以及对手的意图和社会政治需要，潜在地预测对手未来的行动（见5.1节）。

这种上下文使组织能够采取更多的战略缓解措施。例如，该组织现在可以对包含有价值信息的资产进行特定对手的检测和缓解，开展长期的缓解行动（如[11]中所述），识别并与共享威胁空间（见5.1.3节）中的合作伙伴沟通，制定联合缓解计划，共享非技术指标等等。

7 支点分析

支点是提取数据元素并与数据源一起利用该元素发现其他相关元素的分析技术。因此支点是假设检验的基本分析任务。入侵事件的每个元素都会生成自己的假设，这些假设需要证据来加强、削弱或改变。支点是用于发现相关元素（证据）的任务，这些元素用于假设，也生成了新的假设本身。支点的成功依赖于分析人员理解元素之间的关系以及它们成功利用数据元素和数据源的能力（例如，如果我将这些信息与此数据源结合在一起，那么我可以找到这些信息...）。

钻石模型从根本上支持支点分析是其最强大的特征之一。事实上，钻石模型最初是在探索轴心场景后发现的。核心特征被构造成一个“钻石”，其连接边突出显示支点因素，以表明对手行动的其他要素。通过钻石的一个点，分析人员可以发现和揭露其他相关的特征¹¹。

以图5为例：（分析转轴1）受害者在其网络上发现恶意程序，（分析转轴2）恶意程序反向暴露了命令与控制（C2）域，（分析转轴3）域被解析暴露控制恶意程序的IP地址，（分析转轴4）通过检查防火墙日志找出与现在已知的控制恶意程序IP地址进行通信的与受感染的主机同一个网络中的其他受感染的主机，最后（分析转轴5）IP地址注册显示对手的详细信息，提供对手的潜在属性。

7.1 ‘中心’方法

该模型适用于多个集中的入侵分析情报技术概念。这些方法被称为“中心”方法，因为它们以钻石的特定特征为中心，发现新的恶意活动并揭示与其他连接特征和特征本身相关的活动。

7.1.1 以受害者为中心方法

大多数组织通过正常的网络和主机监视、检测和防御行为，正是以受害者为中心的方法。通过这种方法，分析与潜在受害者相关的数据可以发现其他相关（和钻石连接）元素：恶意能力和基础设施。

¹¹ 钻石模型能否成功发现是无法保证的。它只强调可能的，而不是确定的。

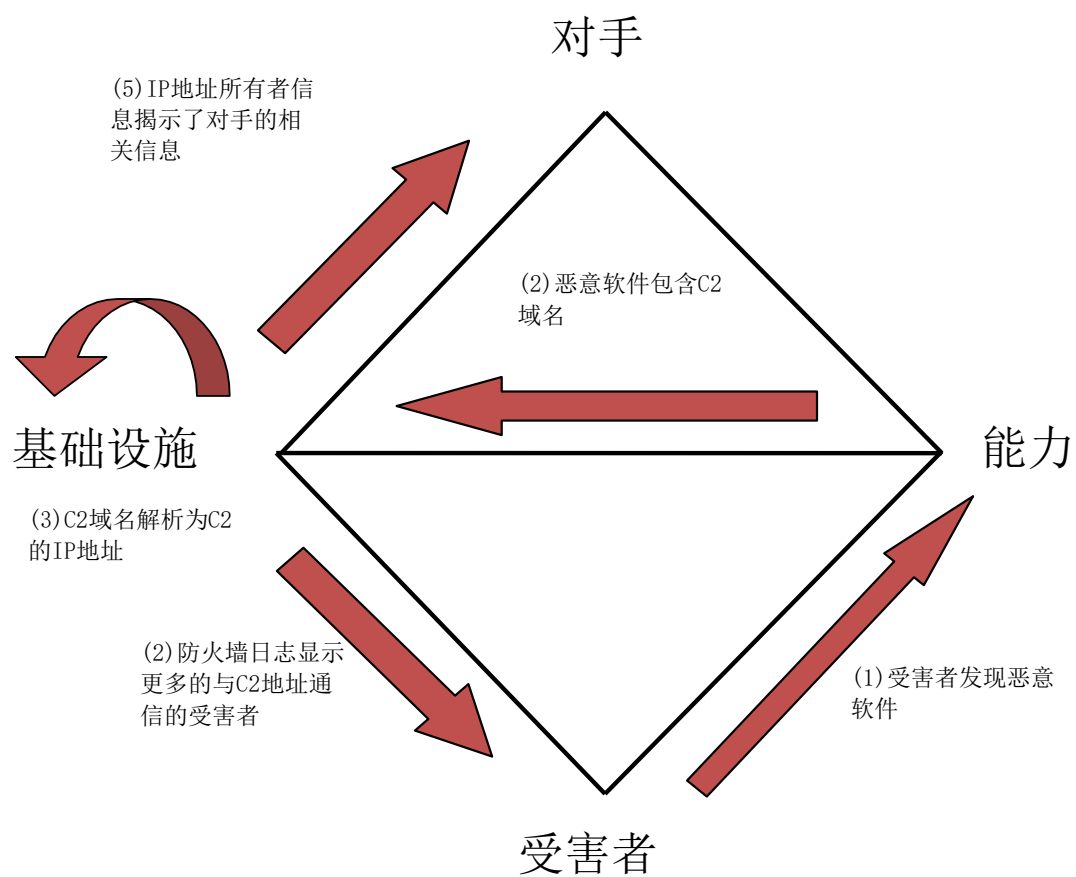


图5：图中即为钻石模型支点分析，作为钻石模型最强大的功能之一，转轴允许分析人员利用特征之间的基本关系（通过特征之间的连接边突出显示），发现恶意活动的新信息。

Honeynet项目就是这种方法的一个很好的例子。他们通过搭建了一个特殊配置主机作为受害者，并邀请对手攻击该主机，揭示其能力和基础设施，然后将其公布用于防御和教学[7]。

以受害者为中心的方法的另一个有趣的例子是，分析人员监控Himalayan为用户提供的服务，这些用户被认为是能力很强的对手的目标[41]。正如钻石模型预测的那样，当对手攻击被监视网络的用户时，这产生了关于恶意能力和基础设施的新信息。有趣的是，这种以受害者为中心的方法与以社会政治为中心的方法（见7.1.5节）相结合，使研究人员能够通过预测受害者、增加成功机会和增强归因置信度来瞄准特定的对手。

7.1.2 以能力为中心方法

以能力为中心的方法利用能力的特点来发现与对手作战相关的其他要素：该能力所针对的受害者、支持该能力的基础设施、实现该能力的技术、其他相关能力的线索和（可能的）对手线索。这种方法的结果在反病毒供应商报告中最常见。

第一个例子，Symantec和CrySyS的分析基于代码中使用的几个共同的特征和技术，提供了从Stuxnet到Duqu的链接，这些特征和技术暗示了一个共同的作者。在这种情况下，这些特性是如此先进，以至于它们可以转到对手的特征，从而推断出潜在的对手应承担的攻击责任。这是一个能够利用社会政治元特征增强归因置信度的对手支点的例子[42]。

第二个例子，Kaspersky对“Red October”的分析提供了一个在多个关键点的以能力为中心的分析中的优秀案例研究。在这里，分析从恶意程序功能开始，并针对技术（HTTP、RC4加密、Zlib压缩）、C2结构和基础设施进行了逆向分析。然后，将该能力与受害者的反病毒检测数据库（从受害者到能力转轴）结合使用，以检测“1000多个不同”的相关文件，这些文件也被逆向，以识别其他基础结构（从能力到基础结构转轴），通过“Sinkholing”¹²技术可以从这些识别到的对手基础设施中识别出全球受害者（从基础设施到受害者转轴）。然后进一步确定每个受害者的社会政治地位（如大使馆、政府、军事、能源），以使读者能够利用网络受害者学推断出潜在的对手，这些对手将具有匹配的社会政治需求（见5.1.2节）[43]。

¹² “Sinkholing”是一种侵略性的防御技术，用于接管对手基础设施的位置，以进行缓解（对手无法再使用其无法控制的东西）和分析（恶意软件和受害者继续与现在由防御者控制的基础设施通信）。

7.1.3 以基础设施为中心方法

以基础设施为中心的方法侧重于对手的恶意基础设施。从这一要素中，可以发现其他相关要素：与基础设施接触的受害者、通过基础设施分发或控制的能力、其他相关基础设施（如恶意域解析的IP地址）和（可能的）对手线索，包括可能直接控制基础设施¹³的人。

Command Five团队在SKHack调查中展示了一种以基础设施为中心的方法[44]。虽然最初的详细信息是从响应过程中发现的恶意程序中收集的，但作者使用已知回调域到IP地址的解析，然后转到whois注册信息，以发现具有相同注册者的许多其他域（从基础设施到对手转轴）。然后，他们成功地绘制了完整的基础设施图，有些基础设施没有在攻击中使用，但很可能由相同的对手控制，可以用于先发制人的防御行动（例如，在作战使用之前阻止网络访问这些域）。对注册域的进一步研究还获得了其他针对不同受害者的攻击中使用的恶意程序的信息，但也可能被同一对手使用（从基础设施到能力转轴）。

7.1.4 以对手为中心方法

可以说，以对手为中心的方法是各种中心方法中最困难的。它包括直接监视对手以发现其基础设施和能力。当然，这可能是最富有成效的方法，但受访问需求的限制。例如，美国联邦调查局（FBI）监测了“Phonemasters”黑客组织的电话呼叫和调制解调器活动，确定了他们行动的全部范围，包括涉及的其他对手人员及其受害者、能力和基础设施[45]。不过，有个说法值得注意，追踪对手太近那么就会付出代价[46]。

7.1.5 以社会政治为中心方法

以社会政治为中心的方法是独一无二的。单独而言，它不会直接产生新的因素或指标，而是利用预期的对手-受害者关系来假设谁可能是受害者，谁什么可能是他们的对手，或者谁可能是敌人和他们的预期受害者。然后，这将导致可以利用以对手为中心或以受害者为中心的方法获取战术细节的要素。

¹³ 分析人员经常通过导出注册信息来分析对手-基础设施之间的关系，但常常被错误的信息所挫败。但是，如果对手坚持使用能被追踪到的公共角色或能在恶意事件之间被发现的信息，那么错误信息（例如在域名注册中）可能会很有用。

实际上从入侵活动和现实政治事件的关联中得出分析结论的情况相当普遍。早在1990年，切斯威克就把入侵活动与1990-91年海湾战争联系起来了。最近，2008年格鲁吉亚的ddos攻击和针对亲藏组织的持续攻击与当时的政治事件有关[47，48]。然而，必须时刻注意相关性不是因果关系。

7.1.6 以技术能力为中心方法

以技术能力为中心的方法使得分析人员针对潜在的误用或异常使用技术来发现先前未知的基础设施和使用此类技术的能力。监测和检测域名系统（DNS）中的异常已经成为一种流行的、富有成效的方法，可以实现以技术为中心的方法来发现新的恶意活动[49，50]。另一些人则研究了骨干网络上数据包头的异常情况[51]。

8 活动线

公理4指出，对手不会在单一事件中针对受害者进行行动，而是在一系列有序阶段的因果事件链中进行行动，通常情况下，每个阶段必须成功执行才能实现其意图¹⁴。活动线是一个有向相序图，其中每个顶点都是一个事件，弧（即有向边）确定事件之间的因果关系。弧被标记为建立因果关系的分析置信度，无论路径是AND（必要的）还是OR（可选的——一个事件有多个潜在路径），无论弧是实际的还是假设的，以及前一个事件提供的信息或资源，这些信息或资源是下一个事件发生所必需的¹⁵。线是垂直组织的，这样每个线都描述了对手针对特定受害者执行的所有因果事件（然而，模型的实现定义了受害者特征），这些事件的共同目标是实现对手的意图。因此，每个线都是特定于一个对手-受害者对的——尽管在许多情况下，活动线在受害者之间可能仅略有不同，仅仅因为对手整合了基础设施、过程和能力以降低成本。

¹⁴ 如4.5.2所述，一组阶段可包括给定活动的非必要阶段，因此并非所有活动都符合可用的完整阶段集。因此，我们说，通常每个阶段都必须成功地执行以实现一个意图，但它不一定适用于所有活动的阶段。

¹⁵ “AND/OR”和“要求/提供”这两个概念都是以前模型中的概念，在不同的缓解战略制定模式中都是有用的。合取攻击路径和析取攻击路径借鉴于Schneier关于攻击树[15]的早期研究，对于可达性、路径优化和其他图分析技术开发缓解策略非常有用。以资源为中心的图的概念来自于[52]，在资源约束缓解策略的开发中是有用的。这并不是说两者必须同时使用，而是提供最大的可能应用不同的技术进行比较，因为在制定缓解战略方面没有任何一种技术被证明是最佳的。然后，可以使用决策支持模型（如ADAM）对这些技术的输出进行比较，以权衡它们的各种风险、成本和收益[13]。为了支持这一点，正如Schneier在[15]中所描述的那样，弧还可以包括权重、优先级或其他量词。

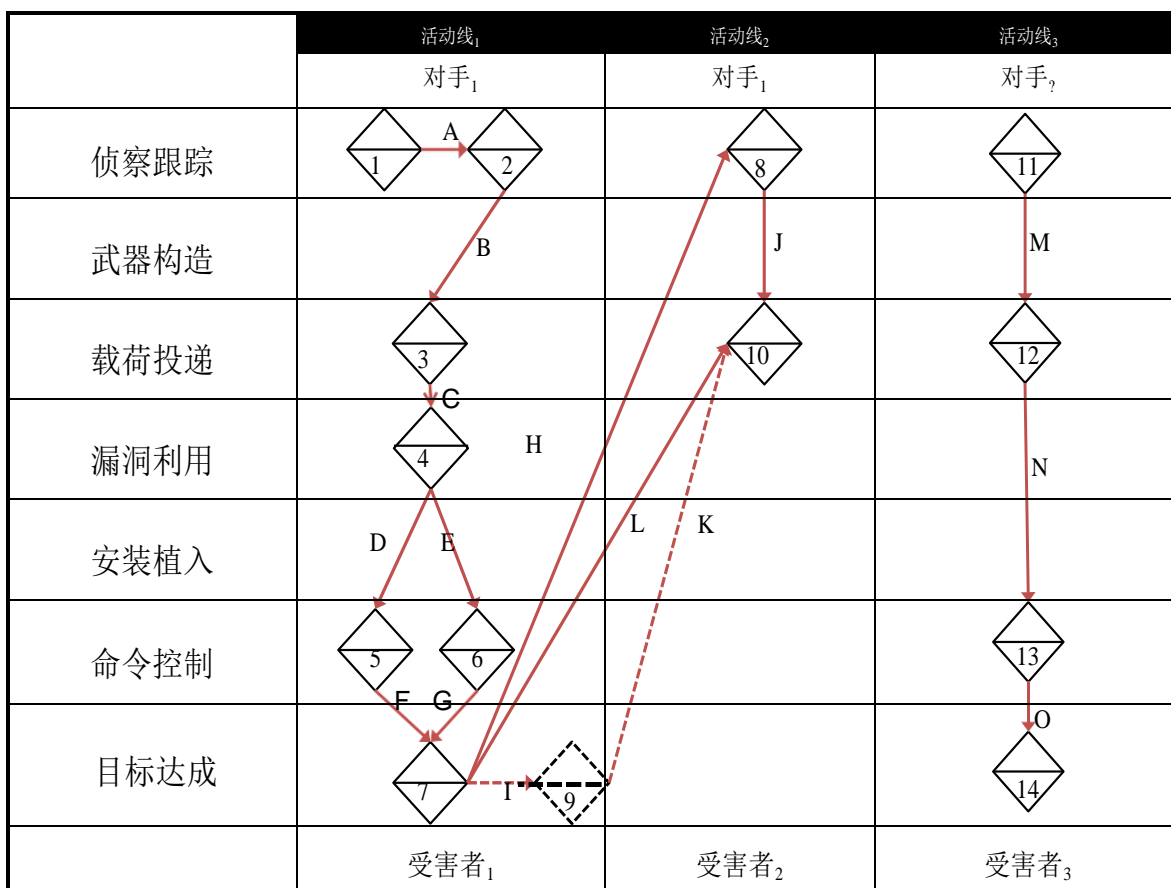


图6: 钻石事件的活动线程的可视化示例, 通过有向弧垂直 (在单个受害者内) 和水平 (在受害者之间) 连接, 表明事件之间的因果关系 (即, 此事件发生是因为此事件, 在此事件之后)。图中实线表示证据支持的信息的事实元素, 虚线表示假设元素。事件描述见表1, 弧描述见表2。

表1：图6的活动线事件描述示例

事件	假设/事实	描述
1	事实	对手公司对受害者 Gadgets 公司进行网络搜索，将其域名 gadgets.com作为结果的一部分进行收集。
2	事实	对手使用新发现的域gadets.com进行新搜索“network administrator gadget.com”发现来自声称是gadget.com网络管理员的用户向论坛发布的邮件，这些用户泄露了他们的邮件地址。
3	事实	对手向gadget.com的网络管理员发送带有木马附件的鱼叉式钓鱼邮件，如事件2所示。
4	事实	gadet.com的一位网络管理员（NA1）打开执行附有漏洞利用程序的恶意附件，进一步执行代码。
5	事实	事件4中被攻陷的主机向IP地址发送HTTP Post请求并注册到控制端，接受HTTP响应。
6	事实	通过对NA1主机上的恶意程序进行逆向工程，可以发现该恶意程序配置了一个额外的IP地址，如果第一台主机不响应，该地址将作为备份。
7	事实	通过向NA1主机发送HTTP响应中的命令与控制消息，恶意程序开始代理TCP连接。
8	事实	通过在NA1的主机上建立的代理，对手1在网络上搜索“有史以来最重要的研究”，并找到受害者“有趣研究”公司。
9	假设	对手1检查NA1的邮件联系人列表中是否有来自“有趣研究”公司的联系人，其中在联系人中发现了“有趣研究”公司的首席研究官。
10	事实	“有趣研究”公司的首席研究官收到来自Gadgets公司的NA1地址发送的鱼叉式钓鱼邮件，该邮件具有与事件3中观察到的相同的攻击载荷。
11	事实	未知的攻击者会扫描易受攻击的Web服务器，包括受害者3。
12	事实	利用先前在事件10中扫描到的漏洞，通过网络将漏洞利用程序部署到受害者3上。
13	事实	被攻陷的服务器也就是受害者3向对手建立一个远程shell连接。
14	事实	对手使用远程shell下载受害者3私有目录中的所有文档。

表2：图6的活动线弧描述示例

弧	置信度	与/或	假设/事实	提供
A	低	与	事实	提供Gadgets公司域名gadgets.com。
B	高	与	事实	提供鱼叉式钓鱼目标：gadgets.com的网络管理员邮件地址。
C	高	与	事实	[无]
D	高	或	事实	[无]
E	高	或	事实	[无]
F	高	与	事实	[无]
G	高	与	事实	[无]
H	中	与	事实	提供从之前受害者到搜索引擎的代理访问。
I	低	与	假设	访问邮件联系人列表。
J	高	与	事实	受害者组织识别。
K	低	与	假设	受害者邮件地址，姓名和角色识别。
L	高	与	事实	鱼叉式钓鱼木马邮件。
M	高	与	事实	提供成功扫描结果的输出，以确定Web服务器是否易受攻击。
N	高	与	事实	[无]
O	高	与	事实	提供建立的远程shell。

垂直相关性 很少出现单条垂直活动线中的所有事件都已知的情况。此外，可能需要努力在需要额外研究、数据收集和分析的活动线内的事件之间建立因果关系。识别知识缺口，用新知识填补这些缺口，并在单个垂直对手受害者活动线内建立因果关系（以及相关的弧标签）的分析过程被称为垂直相关性。通过阶段组织活动线，还可以更容易地识别活动应该发生的但是目前不存在的知识缺口（更多信息请参见8.2节）。

对手通常使用一个操作中获得的支持未来的操作或利用内部信任关系来获得对特定网络的更深入访问，这在渗透测试中被称为跳板攻击和横向利用。因此，因果关系（弧）可以水平跨越一个或多个活动线。此外，如图6所示，阶段可以包含多个事件，弧甚至可以“向后”描述重复的过程，而边则描述在事件之间获得和使用的资源。

水平相关性 分析过程中的因果联系事件之间的垂直线程跨越对手-受害者对，确定线程之间的共同知识差距，并使用知识从一个线程填补在另一个线程中的知识差距被称为水平相关性。这一过程还导致识别受害者之间的共同特征，从而在稍后定义的过程中创建一个活动组（第9节）。

这些活动线形成了一种新的相序攻击图¹⁶，通过对实际事件的观察来预测特定路径的可能性和对手偏好。与传统的攻击图一样，活动线模拟复杂的多阶段行为，这些行为可能利用多个系统和网络漏洞。然而，与试图穷举列出所有可能路径的传统攻击图不同，活动线模拟实际攻击路径的知识以及线内和线之间的相互依赖性。如本节中所定义，活动线的性质允许事件/顶点满足另一个事件（第4.5.6节）的一个或多个资源需求，从而使后续事件发生。此外，每个顶点都是一个事件，事件所提供的信息深度，使图形信息变得丰富，而且本质上更有用。

形式上，我们可以把活动线定义为有向图 AT ，其中 $AT = (V, A)$ 是有序对，其中：

- $|V| \geq 1$ 在线中至少存在一个事件¹⁷。

¹⁶ 攻击图是一个列举了对手渗透计算机网络达到其预期目的可能采取的所有路径。

¹⁷ 虽然公理4确保每个行为至少有两个阶段，但在发现时可能并非所有阶段都是已知的，因此最初只能用一个事件创建活动线程。然后将空阶段和缺失事件视为知识缺⁵⁴。

- AT 是一个有限图。
- V 是所有事件的集合，这些事件被划分成子集合，这样子集合中的所有事件都共享同一个对手和受害者，并进一步被划分为 p 标记的元组，其中 p 是定义的阶段数，每个事件都被放入与其阶段相匹配的元组中。
- A 是一组有序的弧线对，类似定义为 $arc(x, y)$ ，当且仅当对手由于事件 x 成功地执行了事件 y ，并且事件 x 直接位于事件 y 之前。
- 存在多个弧发向任一事件。例如，给定三个事件 x ， y 和 z ，存在从 x 到 y 弧 $arc(x, y)$ 的路径以及从 z 到 y 弧 $arc(z, y)$ 的路径。
- 存在任一事件发出多个弧。例如，给定三个事件 x ， y 和 z ，存在从 x 到 y 弧 $arc(x, y)$ 的路径以及从 x 到 z 弧 $arc(x, z)$ 的路径。
- 从一个节点到另一个节点只能存在一条路径（即每个弧序对在图中是唯一的）。例如，给定两个事件 x 和 y ，从 x 到 y 只能存在一条路径 $arc(x, y)$ 。
- 弧被标记为4元组，置信度、与/或、假设/事实、提供，其中：
 - 置信度：定义 x 和 y 之间因果关系存在时的分析置信度。
 - 与/或：定义从 x 到 y 的路径是否是 y 成功必需（AND），或者该路径是否是从 x 实现 y 的可选替代（OR）路径。
 - 假设/事实：将假设弧（见8.2节假设弧说明）与有证据支持的事实弧区分开来。
 - 提供：定义资源 x 使 y 能够成功匹配资源事件元特征中列出的需求（见4.5.6节）。

8.1 对手过程

总体而言，垂直线和水平链有效地描述了公理4定义的对手端到端过程。事件本身进一步丰富了这一点，其中包含了个人行动的特征（例如，所使用的能力和基础设施、具体的方法、所用的外部资源）。这些共同定义了对手如何实现他们的作案手法。

然而，在许多情况下，对手会表现出对其更广泛过程中某些要素和行为的偏好。这一事实已在犯罪学中得到确认和探索，很可能是由于文化、知识、培训、经验等舒适和熟悉事物对人类的吸引力的结果[53]。在更大的组织中，这些偏好也可能由领导的政策和法令驱动。入侵分析人员通常通过一次行动中的共同因素来确定这些偏好，就像传统的犯罪调查人员通过犯罪现场中的共同证据来确定这些偏好一样。

识别和表达同对手特征和行为的这种能力是非常有用的。通过这种特性，分析人员可以将共享相似过程（请参阅第9节）的类线组合在一起，而不需要为每个事件匹配确切的特征（例如，相同的基础设施IP地址、相同的功能）。钻石模型将此定义为一个对手过程。

例如，图7说明了从图6中的事件2、3、4和6定义的对手过程。这种对抗过程通常被描述为：侦察事件，包括对“网络管理员”的网络搜索，随后（但不一定是立即）发送带有特洛伊木马附件的邮件，然后在本地计算机上进行特定的已知攻击（例如CVE-YYYY-XXX），最后是发送HTTP Post并离开受害者。这条线现在可用于与其他活动线进行匹配，这些活动线显示事件和特性的相同一般顺序。

从形式上讲，对手过程被定义为活动线的子图，其中包含其特征的子集。重要的是，子图可以是“弹性的”，因为它可以被定义为不需要保持它们的严格顺序来有效地匹配另一条线（如图7所示，事件之间的虚线）。换句话说，只有按一般顺序匹配特征才是重要的，但它们之间可以存在其他事件。或者，可以“严格”定义一个对手过程：事件必须在没有干预事件或两者结合的情况下维持其秩序。

8.2 假设分析支持

支持假设产生、记录和检验是活动线的最重要特性之一，它提供了“竞争假设分析”（ACH）[1]等一体化形式分析模型，并遵循了必要的科学严谨性。分析的第一步是确定要解

决的问题。一旦确定了问题，就可以生成、记录和检验假设。



		过程特征
侦察跟踪		网络搜索“网络管理员” [源自事件2]
武器构造		
载荷投递		发送带有木马的邮件 [源自事件3]
漏洞利用		特定的本地漏洞利用（例如CVE-YYYY-XXX） [源自事件4]
安装植入		
命令控制		来自受害者的HTTP POST响应 [源自事件6]
目标达成		

图7：从图6所示的活动线程派生的一个示例对手流程。在这个过程中，事件2、3、4和6的特征被提取到一个子进程中，该子进程可用于与其他线程进行匹配。事件之间的弧线是虚线，说明虽然事件仍然是按阶段排序的，但其他事件可以在它们之间进行干预，而不会扰乱匹配准则。

如前所述，通过将事件置于基于阶段的模型中，可以更容易地识别水平差距。因为公理4指出恶意行为是多阶段的，每个阶段至少应该包含一个事件¹⁸。另一种识别水平差距的方法是使用资源元特征（见4.5.6节）。然后，分析人员可以思考对手是如何满足每个事件所需的资源的，从而产生必要的假设来解决这个问题。

这些假设可以在活动线程中记录下来，并且必须区别于其他事件。这是一个重要的特点，因为大多数分析的失败之一是缺乏书面的假设，而且更严重的是，缺乏对假设和事实的区分。活动线程模型鼓励假设生成和文档记录，从而提高知识的价值和准确性。

一旦假设被记录下来并加以区分，就必须对它们进行提炼和检验。有几种假设检验方法可以用于我们的模型，以确定给定的假设包括其本身和其他假设是否合理。例如，我们可以将证据权重应用于竞争性假设[1]、奥卡姆剃刀定理（在其他条件相同的情况下，更简单的解释通常比复杂的解释更好）¹⁹，稳健性（假设是否与行为的其他方面“相符”），以及对竞争性假设进行归纳和演绎推理的其他形式方法[1]。

例如，图6中的事件10可以在资源元特征中列出以下内容：发送邮件的网络权限、邮件帐户的访问权限、目标邮件地址、附带木马恶意程序的邮件、以及目标的技术水平，构造邮件以绕过过滤器并诱使目标执行恶意程序。事件7（代理访问）或事件8（搜索结果）都不能提供向首席资源官发送邮件所需的资源，特别是他们的邮件地址和角色（目标的技术水平）。因此，事件9被假设为目标信息的来源，使最有吸引力的邮件被发送到正确的目标。

事件9可以通过几种方式进行检验。首先，它是简单和合乎逻辑的，因为它所需要的所有资源都已满足，不需要再假设任何其他事件。其次，它“适合”对手的能力和访问范围。此外，可以收集证据（如主机事件日志）来确定事件是否发生，从而使假设可测量和可检验，以满足科学的严谨性。

这种形式的记录有助于最终实现入侵分析过程中的可重复性，因为其他分析人员可以独立地跟踪行为图，建立自己的假设和结论，并将其与原始的假设和结论进行比较。这一过程建立了对分析性结论的信心，并提高了最终判断的准确性。

¹⁸ 不能保证每个阶段都有一个事件，因为公理4允许非必要的阶段。

¹⁹ 在我们的模型中，可以通过比较一个事件所需的资源数量以及根据当前事件满足了多少资源，而不是假设需要支持更多的事件，从而很容易地对其进行度量。

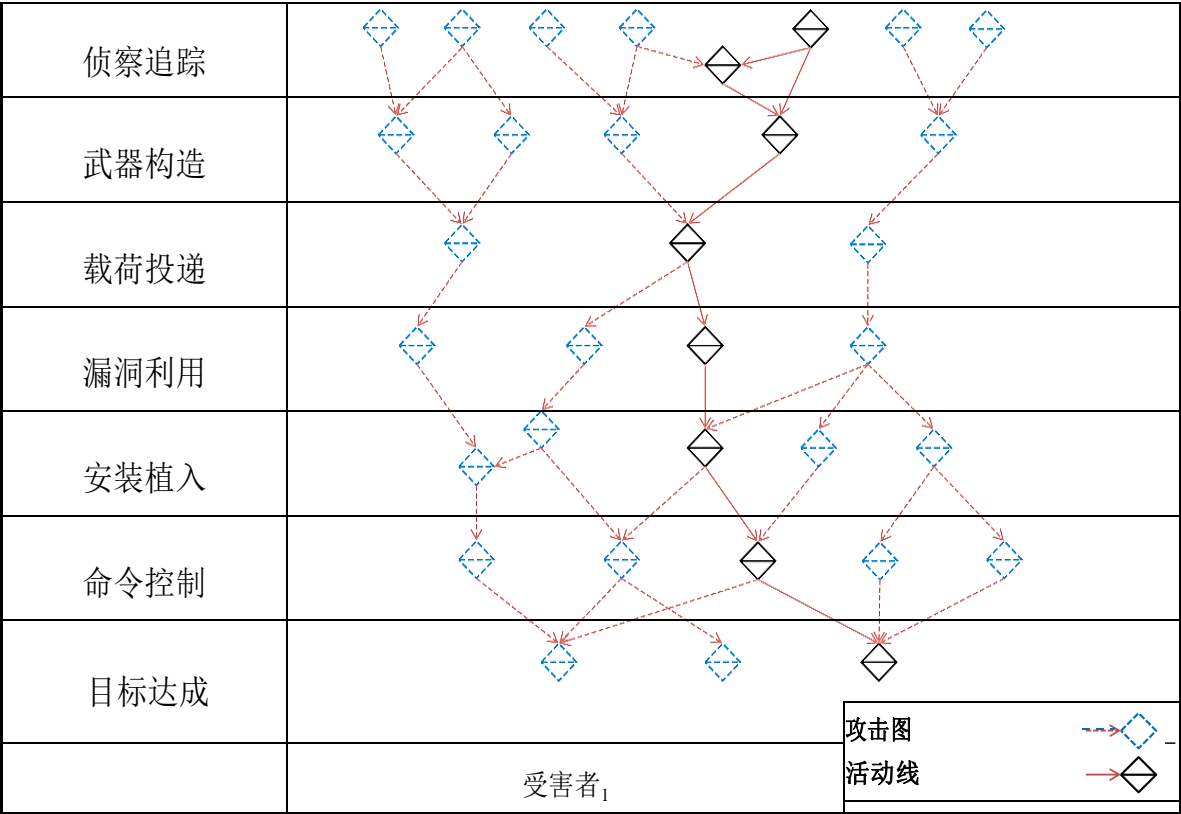


图8：一个行为攻击图示例，展示了实际对手攻击路径的手段和可取的多种假设攻击路径整合。使用行为攻击图可以突出显示对手未来的潜在路径，以及基于当前知识的首选路径。

8.3 行为攻击图

行为攻击图 活动线和传统攻击图并不是互斥的，而是回答了互补的问题。攻击图识别和枚举对手可能采取的路径，而活动线则定义对手了采取的路径。通过将活动线叠加在传统攻击图之上，这些线可以共存。我们将此基于情报信息的攻击图称为行为攻击图。

行为攻击图有几个好处：

- 它保持了攻击图的完整性，使攻击图分析的范围完全可用。

- 它增加了攻击图中包含的信息量，因为每个顶点都是一个富含特征的钻石事件。
- 它增加了攻击图中包含的可视信息量，而几乎没有降低可用性。
- 当实际攻击者选择（和偏好）已知时，它会生成更精确的权重。
- 它突出显示了攻击者的偏好以及其他替代路径。
- 它详尽地（由于攻击图的性质）为博弈场景和缓解策略的制定绘制了备选路径（例如，如果采取这一行动，对手很可能选择其中一条路径）。
- 它通过覆盖与水平相关的攻击线的语料库来进行比较，自然有助于填补任何一个攻击线的知识空白。这样结果就是对正在进行的事件响应调查中更准确、更快地生成和测试假设。

图8.3是行为攻击图的一个示例。该图将已知的对手路径（行为图）与已知的可利用路径（攻击图）区分开来。这很像同时参考渗透测试人员（红队）和脆弱性评估（蓝队）结果，以制定最佳的行动方案²⁰。

最终，活动线和行为攻击图有助于更好地制定缓解策略，因为它们将信息保障和威胁情报结合在一起。它们将已经发生的情况与可能发生的情况结合起来，这样既可以制定应对当前威胁的战略，又可以制定有效应对未来对手行动的响应计划。这种综合规划还可以提高资源利用效率，因为可以在制定缓解策略的同时应对当前威胁以及未来威胁。

9 活动组

活动组 活动组是一组钻石事件和活动线，这些事件和活动线根据它们的特征或过程的相似性关联，并根据置信度进行加权。

活动组有两个目的：（1）提供需要广泛行为知识的分析问题的框架。（2）制定具有比活动线更广泛预期效果的缓解策略。活动组从两个方面区别于活动线：（1）活动组同时包含事件和活动线。（2）活动组中的事件和线程由相似的特性和行为关联，而不是因果关系（与

²⁰ 目前还未探讨行为攻击图方法是否是一种将实际的红蓝队结果整合到一起进行分析的有用方法，但这是一个有趣的问题，留给以后的工作来解决。

活动线的情况一样）。

分析人员通常按照传统方式使用相似的基础设施和能力形成活动组来识别事件和线索背后的共同对手。但是，活动组这个概念本质上是灵活的，并且可以扩展到包括任何基于相似性的分组，满足大量的分析和操作需求。所需的分析或操作结果决定了所使用的相关的实现和类型（即分组函数）。此外，活动组不是静态的，正如对手不是静态的一样。活动组必须随着时间的推移不断发展和变化，以吸收对手的新知识，包括其需求和行动的方面变化²¹。

围绕活动组的流程有六个不同的步骤：

步骤1：分析问题 通过分组来解决特定的分析问题。

步骤2：特征选择 选择用于形成分类和聚类基础的事件特征和对手过程。

步骤3：创建 活动组是由一系列事件和活动线创建的。

步骤4：增长 当新事件加入模型时会被分进活动组。

步骤5：分析 对活动组进行分析解决所定义的分析问题。

步骤6：重定义 需要不时地重新定义活动组保持其准确性。

在形式上，我们将活动组 AG 定义为一组事件和活动线，这些线在特征或对手过程中具有一个或多个相似之处：

²¹ 聚类和分类虽然是强有力的工具，应该鼓励使用它们增强分析力，但它们也并非没有缺陷。聚类和分类方案有许多有据可查的问题。有些问题尤其令人担忧，因为对手几乎在每个数据包中都积极地进行否认和欺骗，以逃避检测和分析。最受关注的点是过度拟合，其中分析人员或机器将不相关的信息包含到集群中。产生这个错误的主要原因是组定义差（弱特征向量）。尤其是在入侵分析的情况下，它显示了两个行动重叠特性（共享一个公共功能，使用共享主机空间）。错误传播进一步加剧了这种情况：一旦将一个不相关的事件包含在组定义中，那么它将用于将来的相关性，增加非相关事件的数量，从而使初始错误更严重。目前存在几种检测和防止过拟合[54]的技术。将这些技术应用于入侵分析的比较超出了本次的范围，留给以后的工作来讨论。不过，这是一个值得注意的问题。

$$AG = \{et_1, et_2, \dots, et_n\}$$

其中：

- $n \geq 1$ ，活动组中必须至少有一个元素。
- et_n 是指第8节中定义的单一事件或活动线。
- AG 中的所有事件或进程都有一个或多个相似之处，满足用于划分事件和线程的活动组创建函数（在第9.3节有定义）。

9.1 步骤1：分析问题

这些问题通常需要基于一些共同的特性（特征向量）进行演绎和推理。这些问题通常是截然不同的，需要为每个问题使用不同的特征向量²²。例如，按可能的对手对事件和线进行分组的特征向量并不总是足以对事件进行分组，从而发现常见的恶意程序作者/开发人员。必须首先定义分析问题

因此，我们将分析问题PR定义为一种入侵分析问题的陈述，它需要聚类和分类（分组）来解决部分或全部的问题。

活动组支持的一些分析问题示例：

- **趋势：**对手的行动如何随时间而变化，以及通过当前的向量如何预测未来的变化？
- **意图推测：**对手的目的是什么？
- **归因推断：**哪些事件和活动线可能由同一对手产生？
- **对手能力和基础设施：**对手完整的观察能力和基础设施是什么？

²² 但并不排除两个或更多问题共享一个共同特征向量的可能性。

- **跨能力识别**：哪些能力是被多个对手使用？
- **对手行动水平差距识别**：在对手的行动过程中组织的水平差距是什么？
- **自动缓解建议**：当检测到事件背后的对手时可以/应该采取什么行动²³？
- **常见能力开发推测**：哪些能力显示了共同作者/开发人员的证据？
- **重心识别**：哪些资源和流程对活动和/或战役最为常见和关键？

9.2 步骤2：特征选择

钻石事件和活动线以两种互补的方式相关联和分组：（1）使用各种事件核心特征，元特征和子特征（例如，基础设施，能力），（2）先前定义的对手过程（见第8节）作为活动组的子图。为此，我们选择了一些特征来填充一个特征向量，该特征向量定义用于对事件和活动线程分组的元素。

重要的是，特征向量可以非常具体，允许分析人员通过包括特定的可观察对象（例如，IP地址，域名，恶意程序）来定义特定的感兴趣活动，从而形成两个组：属于行动一部分的事件和线程，以及不属于该行动的事件和活动线。

此外，包含在特征向量中的流程是一个强大的概念，它不仅通过可观察的方式，而且还通过特定的手段（无论具体的基础设施或能力如何）来支持比较活动。这对于那些可能经常改变基础设施和能力（最常见的可观察性），但却保持半稳定过程的对手来说尤其有效。

特征空间包括事件的所有核心特征和元特征（如基础设施、能力、受害者、结果）（见4节）以及任何定义的对手过程（见8节）。从特征空间选择和/或创建²⁴最相关和最佳的特征，定义特征向量。最后，这些特性中的每一个都可以与一个权重相结合，以确定其在定义组时的相对重要性。有几种著名的技术可以用来选择（并可能创建）最相关和最佳的特征[55]。关

²³ 活动组支持实时智能驱动的网络防御。当实时检测到事件时，应用机器学习分类技术对已知活动组的事件进行分类和关联。给定一组预先确定的条件（例如，如果事件E被归类为具有大于80%置信度的活动组X），系统可以向网络防御机制提出建议，将抑制技术应用于活动。这样，即使对手在不要求防御者预测变化的情况下改变了部分行动，也可以实时抑制对手的行动。

²⁴ 特征创建（从现有特征创建新特征）是因为分析人员通常使用一个函数来比较基于入侵活动的特征。例如，如果原始功能列表中不存在IP，则解析域名的IP地址可能是提取的特征，现在允许引用同一域或其关联IP的事件进行关联。

于活动分组的最佳钻石模型特征选择/创建的进一步讨论，将作为未来研究的一个领域，并且也将是具体的实现。

表3: 示例活动组定义步骤1和2

分析问题	哪些事件和线程可能由使用某个过程 ($Process_1$) 的同一对手执行? (如属性)
特征空间	$Infrastructure_{IP}, Infrastructure_{Domain},$ $Capability_{MD5}, Victim_{IP}, Victim_{Organization},$ $Methodology, Process_1, Process_2, Process_3$
特征向量结果	所有事件和线程将按基础设施IP, 能力MD5哈希和定义的攻击过程 $Process_1$ 中的相似性进行分组。

我们将特征空间 FS 定义为事件以及任何和所有对手过程的所有核心特征, 元特征和子特征的集合。

此外, 我们定义特征向量 FV_{PR} 以解决分析问题, 如下:

$$FV_{PR} = \langle \langle f_1, w_1 \rangle, \langle f_2, w_2 \rangle, \dots, \langle f_n, w_n \rangle \rangle$$

- $n \geq 1$, 特征向量中必须至少有一个元素。
- $f_n \in FS$, 特征向量中的每个特征都必须存在于特征空间中。
- $FV \subset FS$, 特征向量是特征空间的子集。
- f_n 是对事件和线程进行分组以解决分析问题 PR 的必要元素。
- $w_{f_n} \in \mathbb{R}$ 且 $0 < w_{f_n} \leq 1$, 权重是一个实数, 它描述了 f_n 对所有其他 f 的相对重要性, 因此 $w = 1$ 是一个具有最重要性的特征²⁵。

²⁵ 在特征向量中应该没有给定权重为零的特征, 因为这表明该特征没有重要性。在这种情况下, 该特征不应包含在特征向量中

9.3 步骤3：创建

分析人员最初通过认知聚类过程创建活动组：分析人员将事件的特征与其他特征（例如它们的特征向量）进行比较，并使用一些定义相似性的函数将事件划分为不同的组（即集合），并确定事件所属的组的相关置信度。形式上，这些组可以成为应用机器学习技术的类，例如活动组变化中的分类（见9.4节）。

预计一个组织将定义多个分析问题（通过步骤1）。因此，每个钻石模型实例可能有多个活动组创建函数。例如，根据明显相同的行为者-对手对事件进行分组（例如属性分组）和根据受害者脆弱性分组事件（例如，对最有可能的利用路径进行分组）是不同的分析问题，需要不同的函数。此外，有些问题可能需要集群函数将每个事件和活动线放入一个组中，其他事件和活动线可能允许出现异常值（即，不属于任何组的事件和线程）。

活动组的创建是解决特定分析问题的事件和线程关联的通用集群问题。聚类函数依赖于先验信息，如分析问题/目标和特定的特征向量，这些信息对于钻石模型的任何给定应用都是唯一的。因此，很可能没有一个活动组创建函数可以解决所有的入侵分析问题。我们期望进一步研究定义函数优化特定入侵分析问题的聚类，例如针对对手属性的优化事件聚类。

在形式上，我们定义了一个活动组创建函数 AGC ：

$$AGC(PR, FV_{PR}, ET) \rightarrow AGS$$

$$AGS = \{AG_1, AG_2, \dots, AG_n\}$$

- PR 是一个满足函数定义的解析问题。
- FV_{PR} 是满足解析问题 PR 的特征向量。
- ET 是要分组的所有事件和线程的集合。
- AGC 基于特征向量 FV_{PR} 将事件/线程集 ET 的所有元素划分为一组 n 个活动组 AGS 。
- 由 AGC 组成的函数可以使用 FV_{PR} 中定义的特征和过程在集合 ET 内的所有元素上

进行操作²⁶。

- AGS 是一组活动组，其中每个活动组 AG_n 都满足活动组的定义。
- 创建函数可能不会建立任何组，因为不存在相似之处，因此 $n \geq 0$ 。

9.3.1 活动组创建示例

图9说明了如何使用带有异常值的严格划分来定义活动组创建函数（ AGC ），以回答上一节特征向量部分（见9.2节）中提出的基于可能使用相同基础设施IP，能力和三步过程的问题。图中显示了一组17个事件和线程，这些事件和活动线根据我们的函数和特征向量进行分组，以创建三个组、两个错误事件和一条错误活动线，这些事件和活动线不符合该函数的标准，并且保持未分组状态。我们的函数将两条活动线分到了活动组3中。

对于我们的示例，函数中表示的逻辑表明，包含流程 $A \rightarrow B \rightarrow C$ 的任何线程都是一个活动组。如果每个线程中至少有一个事件（不一定在指定的进程中）共享一个基础设施IP和MD5散列，且至少具有中等可信度，那么两个或多个进程匹配线程将在同一个活动组中关联。既然数据是被组织起来回答分析性问题，可以对组进行增长（见9.4节）和分析（见9.5节），从而提供可能回答问题的洞察力。

9.4 步骤4：增长

分析人员通过模仿置信度加权概率分类的认知模式识别过程不断增长活动组：分析人员发现恶意事件，根据特征相似性及其置信度将事件与所有其他已知事件进行比较，并将事件与最相似的组（类）（或者，如果信任不满足其阈值，则放弃关联）。当事件和线程被发现、检测或接收到组中时，此操作会不断增加活动组。

图10展示了活动组的增长：当事件和线程被发现、检测或接收时，根据定义的特征向量将它们分类到不同的活动组中。在本例中，符合条件的线程被分类到活动组2中，而其他线程和事件被分类为异常值²⁷。

²⁶ 活动组创建函数中的操作非常广泛，不需要使用特征向量的所有元素。

²⁷ 但是，正如前面在步骤3，活动组创建（见9.3节）中所描述的，聚类函数是由分析人员的需求和要解决的特定分析问题定义的，因此可以使用备选的集群类型，备选的集群类型可能不使用离群值，而是将每个线程和事件放入一个组中。

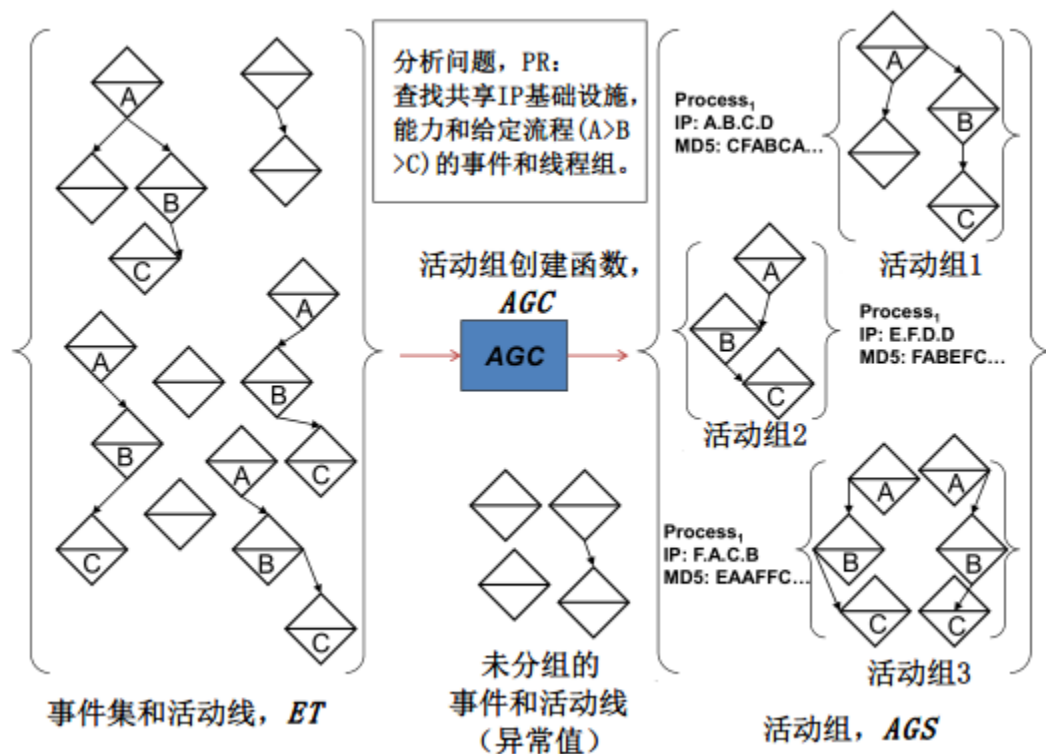


图9: 活动组的创建, 使得一组事件和线程基于定义的特征向量进行集群: 一个对手进程 ($A \rightarrow B \rightarrow C$)、一个匹配能力MD5散列和基础设施IP地址。基于该特征向量和活动组创建函数 (AGC), 将17个事件和线程聚为三个组, 其中有两个事件和一个线程不满足分组标准, 并将其归类为异常值。

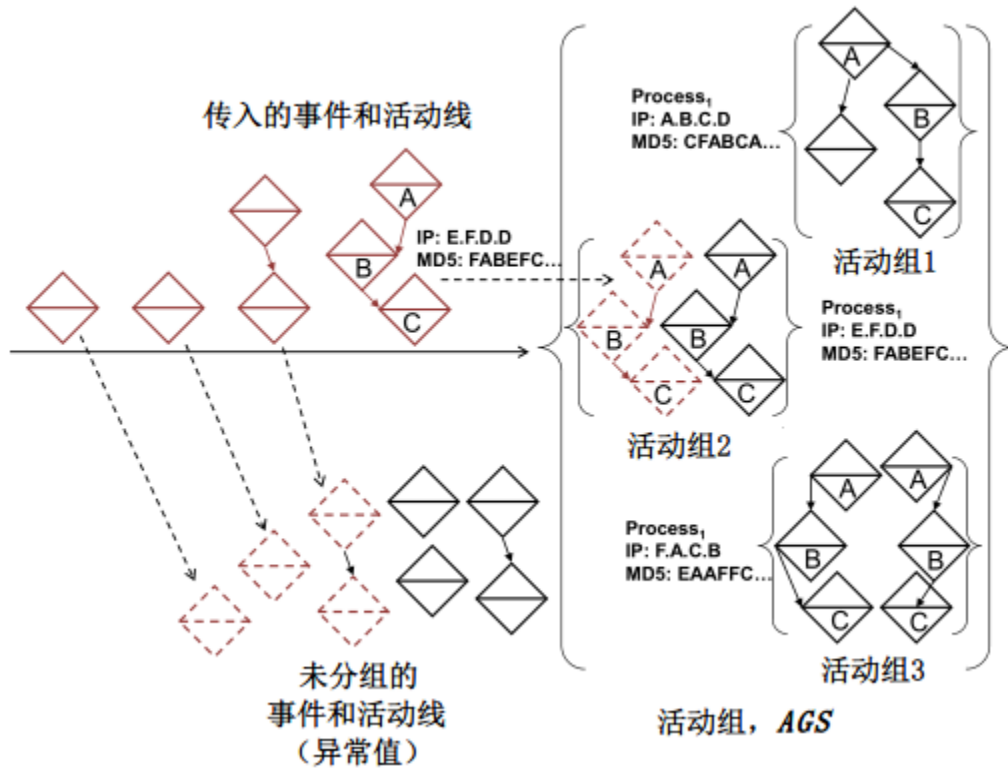


图10: 图10说明步骤4, 活动组增长, 事件和线程被发现、检测或接收, 并根据先前定义的特征向量不断分类为现有的活动组。在这种情况下, 那些没有成功匹配条件的事件和线程是异常值, 而不是分组的。

9.5 步骤5：分析

一旦定义了活动组并将事件和线程聚集在组中，就可以对其进行分析，以解决正在处理的特定分析问题。这通常需要应用钻石模型以外的工具和技巧。例如，通过我们在图9中的示例，分析人员现在可能会检查这些组中的每个组，以识别出需要解决的新的分析问题的不同和相似之处。这甚至可能导致重新检查需要重新定义的特征选择和分组函数（下一步）。

然而，分析人员现在可以分析更大范围内的入侵事件和活动线，包括：潜在地暴露更远范围的对对手活动，识别看似不同的事件之间的相似之处，收集观察到的对手能力和基础设施的完整列表，根据受害者集（例如：网络受害者心理学）推断对手属性，以及许多其他问题。

9.6 步骤6：重定义

与所有基于聚类和分类的函数一样，活动组也面临着各种各样的研究挑战。其中一个挑战是假设分析人员能够准确地描述用于聚类的特征向量和函数，或者他们对聚类的想法从一开始就是正确的。另一个挑战是过度拟合和错误传播：在这种情况下，分析人员或系统错误地将事件关联到一个正在传播的组，并可能随着时间的推移而放大该错误。因此，随着时间的推移，正常的情况下活动组需要检查、异常检测和重新定义（重新聚类）来发现和纠正错误。此外，在重新定义阶段，可以（也应该）考虑特征向量以及相关的权重和算法的更改，以确保底层误差得到纠正。这通常是通过人工发现的证据来完成的，这些证据表明发生了分类错误情况，需要重新聚类。

9.7 活动组族

活动组与恶意活动背后的企业一样多种多样。因此，识别和检测数百万个事件会很容易地过滤到大量活动组中，其中一些活动组在更高级别上进行交互。

因此，有时有必要建立一个群体层次结构，对事件背后日益复杂的组织进行建模，以解决更高层次的问题，并制定更具战略性的缓解措施²⁸。

与活动组非常相似，活动组族是一组具有共同特征的活动组，但一个族中的组的共同特

²⁸ 恶意活动背后的这种组织的证据在“语音大师”案[45]中很明显，布伦纳在[56]中有据理力争地指出现有的等级有组织犯罪模型必然会进入网络空间，因为它们是所有犯罪企业最有效的方法 - 并且网络空间将是犯罪活动的自然延伸，尤其是最大的犯罪活动。但是，我们的模型并不局限于网络犯罪，而是可扩展到任何有组织的企业，这些企业进行大量必要的恶意网络活动。

征可能是非技术性的。例如，在有组织犯罪的情况下，多个活动组由于共同的资金和任务元素的原因被分组到同一个族中，因此多个活动组（每个活动组都是单独跟踪和分析的）被分组在一个族中。这使高阶元素（如本例中的犯罪头目）的识别、组织和缓解策略的制定变得更容易处理也更有效。

出于分析方法的目的，活动组族被视为与活动组相同的6步过程。必须定义、创建、增长、分析和重定义它们。它们还具有特征向量和创建函数（除了跨整个组的特征而不对单个事件或活动线用于聚类 and 分类的创建函数），这些术语以及相关的函数、特征和过程不需要重定义，因为之前已经对它们进行了全面讨论-只要对它们稍作了修改便可以适配跨活动组的特征和过程。

形式上，我们将活动组族定义为：

$$AGF = \{AG_1, AG_2, \dots, AG_n\}$$

其中：

- $n \geq 1$ ，一个活动组族必须至少包含一个活动组。
- AG_n 满足活动组定义。
- AGF 是一组具有一个或多个相似性的活动组。
- AGF 满足一个特殊的解析问题。
- AGF 是比较活动组的创建函数和特征向量的结果。

10 规划和博弈

从缓解的角度来看，可以采取多种措施，然而，决定采取何种最佳行动方案来对抗对手是一项挑战。采取行动需要花费维护者的资金和（或）时间，并且也只能期望防御行动会对对手的攻击起到缓解作用。

我们的模型提供了对手组件之间依赖关系的理解。为了使对手的努力取得成功，必须提供完整的活动线，以此在意图和结果之间创建一条通道。该模型通过确定攻击者需要替换/修复/重新实施哪些组件来帮助理解防御者行为将如何影响对手能力。

另外，防御者应该选择成本较低但是对于对手来说攻击成本很高的行动。显然，从战术或战略缓解的角度来看，相反的情况（即，让防御者付出更多的代价，让对手付出更少的代价）是不可取的。如果可能的话，应该避免让防御者付出更多代价（尤其是显著增加）的行动。对手的成本可以表示为补偿拥有功能平台所需的能力和基础设施的成本（金钱、资源、时间）。对手成本有多个组成部分，包括开发时间、基础设施建设成本/时间、再培训时间和成本、机会成本和因准备不足而产生的成本。防御者成本也有多个组成部分，如金钱、时间，以及需要解决的法律和道德风险[13]。

钻石模型是一个基本概念，有助于构建和加强分析，以实现其最终目标：缓解。该模式没有规定缓解战略或行动方针。它们独立于模型而存在。相反，它支持多种形式的决策。以下是关于该模型对几个流行决策框架各方面的适用性讨论：

联合作战情报环境准备（JIPOE） 美国国防部军事规划条令联合作战情报环境准备（JIPOE）[12]是一份众所周知且经常被引用的指南，它建立了一个利用情报制定行动方案的过程。该指南中提到，如果仅仅是废除对手的基础设施和能力，那么相当于打了一场败仗。相反，它提出了一种还包括如何确定对手的资源、重心、反应和行动方针的综合方法。这一方法可以对抗对手的维护和重建能力，一旦这些能力和基础设施得到缓解，那么便确定了最佳的缓解区域。我们的模型支持这样的规划，因为它：

- 通过缺少活动线中的事件特征和相位差来帮助识别情报和信息差距（JIPOE步骤1，要素6）。

- 支持对手模型演变（JIPOE步骤3，要素1）。
- 识别对手的基础设施和能力，重点关注资源（JIPOE步骤3，要素3）。
- 通过活动线和活动组分析识别对手重心（JIPOE步骤3，要素4）。
- 通过活动线分析、受害者分析和活动组识别对手的目标和最终状态（JIPOE步骤4，要素1）。
- 通过活动攻击图分析确定可能的对手行动路线，在该图中可以识别潜在的和首选的攻击路径（JIPOE步骤4，要素2和3）。

杀伤链分析 钻石模型和杀伤链分析是高度互补的。杀伤链分析允许分析人员“锁定目标并与对手接触以产生预期效果”。钻石模型允许分析师发展间谍情报技术，理解为了建立和组织执行杀伤链分析所必需的知识。本文描述了将这两种方法结合起来的两种方法：

- 一旦分析人员发现了一个活动线程，就可以使用杀伤链行动矩阵来识别活动线中每个事件的行动过程。如图11所示，图6中为活动线1和活动线2标识了每个杀伤链阶段的行动过程。钻石模型的强大之处在于，行动过程可以设计成跨越多个受害者，跨越对手的活动，从而在进一步降低对手能力的同时，使这些行动变得更加有力。
- 由同一可能对手聚在一起的活动组（即按属性聚集），通过分析组内事件中最大的共同特征集，可以提供杀伤链所需的关键活动指标，以集中行动路线并确定其优先级。

漏洞覆盖 常见的信息保障做法是分析系统（或网络）是否存在漏洞，根据组织的具体关注点（如资产价值和成本）对这些漏洞进行排名，然后采取缓解措施修复这些漏洞。通过生成活动攻击图，图中的哪条路径通过缓解修改（从而拒绝由对手使用的路径）的正常过程由对手的偏好和潜力决定。通过使用我们的模型，传统的信息保障决策不再是假设潜在的对手及其路径，而是将实际的攻击路径注入到图中，并预测出对手的偏好和潜力。这种方法提供了更完整的保护，并增加了对手的成本，因为他们现在必须在现有能力基础之外进行开发、训练和操作。

	检测	拒绝	干扰	弱化	欺瞒	破坏
侦察追踪	网络分析	阻止论坛使用政策			创建假日志	
武器构造						
载荷投递	网络入侵检测系统，用户教育	邮件病毒扫描		邮件队列	过滤但使用外出邮件进行响应	
漏洞利用	基于主机入侵检测系统	补丁	数据执行保护			
安装植入						
命令控制	网络入侵检测系统	HTTP白名单	网络入侵防护系统	HTTP节流		
目标达成	代理检测	防火墙访问控制表	网络入侵防护系统	HTTP节流	蜜罐	

图11：图6中线1和线2派生的杀伤链行动矩阵。确定了每种类型的缓解措施（如干扰、弱化、拒绝），以对抗各阶段对手事件的有效性。这一矩阵格式和过程在[11]中被描述为一种确定对抗对手运动的缓解行动方针的方法，说明了将钻石模型和杀伤链分析结合起来的能力。

博弈 任何有效缓解策略制定的一部分都是与对手博弈，以预测他们的下一步行动。这样，防御者既可以反击当前的攻击活动，也可以为将来的攻击活动做准备，从而达到对抗对手的目的。同时为这两种需求进行规划可以实现满足这两种需求的经济决策。我们的模型可以在很多方面更准确预测对手对环境压力（如防御行动，修补漏洞）的反应：

- 它能够围绕人类决策进行更高层次的博弈，因为社会政治和意图元特征是一个整体（例如，对手的需求和愿望是什么？如何影响或对抗这些？）。
- 通过活动线和活动组，可以通过网络受害者学（见5.1.2节）和其他方式进行归因推断。
- 假设检验的基本支持以此具有更完整的博弈场景能够提高博弈的价值和其结果的准确性。

11 未来工作

目前的钻石模型还只是认知水平的、高度依赖手工的，正如其名它只是一个为了准确捕获入侵分析过程而被研究和提炼定义出的模型。不过，我们是实用主义派，如果钻石模型实现了其自动化和效率化，那么将会发挥更大的作用。因此，希望我们已经提供了足够多的见解和相关的努力，能够推动今后沿着这些方向开展的工作。

其中一个非常宝贵的自动化设计就是将钻石模型集成到分析工具中，这些工具既可以自动从网络传感器处获取情报，也可以从其他组织（特别是共享威胁空间中的组织）获得外部报告（见5.1.3节）。我们预计入侵分析人员仍然需要输入新的情报并检查自动化的反馈。这将需要对可用性进行研究，以增强而不是阻碍分析工作流程。此外，还可以实现自动化和形式化的假设生成和测试，提供对分析结论的即时评估。

为了实现这一点，必须有一个协议来共享上下文指标和威胁情报，以快速整合来自所有这些来源的信息。我们将钻石模型视为实现这一目标的基础，改进新的或现有的协议和形式语言（如[30, 27, 22, 24, 32]）以使它们更具上下文性和关系性。当然这也可能需要进一步完善分类法。钻石模型本身提供了将特征和子特征定义为永无止境的信息层次的能力。但是，模型的不同实现可能会在其定义中发生冲突。因此，使用分类基本原理进一步细化每个特征

和子特征的子模型至关重要[25]。

还有一些其他方面的工作需要未来的努力，例如：

- 针对特定分析问题的特征向量和聚类/分类算法的定义。
- 将渗透测试和漏洞评估输出潜在地集成到活动攻击图中。
- 聚类/分类过程中防止入侵分析事件过拟合的方法。
- 将事件子特性作为分类法进行彻底的检查和定义。
- 评估变量和方面以确定攻击持久程度。
- 更全面地了解社会政治领域及其在缓解决策中的作用，包括对抗对手的需求和愿望。

最后，该模型的目的是实现更有效准确的分析，最终使规划、战略和决策都能够起到保护网络的作用。虽然我们展示了如何将钻石模型与其他几个规划框架结合起来使用，但是每个框架本身都可以看作是一种工作，况且还需要考虑许多其他模型。

例如，在一个共同进化的捕食者-猎物环境中，利用遗传算法扩展[14]工作，并将活动线作为一组染色体来对待，可能会产生更有效和更具创造性的缓解策略。这种方法已经展示出了良好前景，活动线模型是一种表现良好的遗传算法，从而为这一概念提供了依据。

12 结论

本文提出了入侵分析的钻石模型。它始于所有入侵行为的原子元素，事件以及以钻石形状组织的核心特征（对手，受害者，基础设施和能力）。事件进一步使用子特征和元特征进行了改进，允许它包含和关联恶意事件的所有方面。从事件中，我们得到了几种以特征为中心的方法，以帮助分类现有的分析技巧以及开发新的技巧。该模型对恶意活动提出了新理解，比方说对手和受害者之间的社会政治关系的重要性以及攻击持久程度。

此外，钻石模型将入侵活动的本质看为一组因果事件，这些事件与记录对手端到端过程的活动线相关。重要的是，这些线还进一步增强了攻击图，为传统的信息保障创建了一种新

的智能驱动方法，称为活动攻击图，活动攻击图同时考虑到实际的对手攻击以及潜在的和首选的路径。然后将这些活动线和事件合并到活动组中，以解决更广泛的分析问题，并制定更多战略性缓解措施。最后，活动组可以分层组织为家族，从而能更好地模拟复杂的对手组织。

该模型还被证明与多种缓解规划和决策模型具有高度互补性，包括联合作战情报环境准备、杀伤链、传统信息保障脆弱性覆盖和对手博弈方法。

长期以来，入侵分析一直被认为是一门需要学习和实践的艺术，而不是一门需要研究和提炼的科学。这方面的证据无处不在：从注重分析结果而不是过程和原理，到通过故事和案例研究传播知识。长期以来，仅将入侵分析作为一门艺术来对待，改进和理解工作一直在拖延，进一步减缓了依赖于高效、有效和准确分析的威胁缓解措施的发展。在不知情的情况下，分析师人员使用钻石模型已有数十年，但缺乏完整的框架来理解、改进和集中精力。

是时候认识到这门学科既是一门艺术也是一门科学了。钻石模型解决了将入侵分析的艺术和科学结合在一起的这一挑战。钻石模型准确地捕获和组织了基础和基本概念，这些概念支撑了入侵分析人员的所有工作，以及如何综合入侵分析并将其用于缓解和网络防御。它既是一种非正式的认知分析支持，又是一种将数学和计算概念应用于入侵分析的正式框架。不过，其最大的贡献在于，它最终将科学的严谨性、测量原理、可校验性原则和可重复性原则应用到入侵分析领域中，从而使入侵分析变得更加有效、高效和准确，快速高效地缓解攻击，最终击败我们的对手。

参考文献

- [1] Richards J. Heuer Jr. Psychology of Intelligence Analysis. Central Intelligence Agency, 1999.
- [2] Chris Sanders. The 10 commandments of intrusion analysis. [ONLINE] <http://chrissanders.org/2011/01/the-10-commandments-of-intrusion-analysis/>, January 2011.
- [3] Leo Obrsta, Penny Chaseb, and Richard Markeloffa. Developing an ontology of the cyber security domain. In Paulo C. G. Costa and Kathryn B. Laskey, editors, Proceedings of Semantic Technologies for Intelligence, Defense, and Security (STIDS) 2012, pages 49-56, October 2012. [ONLINE] <http://ceur-ws.org/Vol-966/>.
- [4] Clifford Stoll. Stalking the wily hacker. Communications of the ACM, 31 (5) :484-497, May 1988.
- [5] Steve Bellovin. There be dragons. In 3rd Usenix UNIX Security Symposium, Baltimore, MD, USA, September 1992.
- [6] Bill Cheswick. An evening with Berferd. In Firewalls & Internet Security, chapter 10. Addison-Wesley, Reading, MA, USA, 1994.

- [7] Lance Spitzer. The honeynet project: Trapping the hackers. *Security & Privacy*, page 15, April 2003.
- [8] Stephen Northcutt, Mark Cooper, Matt Fearnow, and Karen Frederick. *Intrusion Signatures and Analysis*. New Riders Publishing, Indianapolis, IN, USA, 2001.
- [9] SANS. [ONLINE] <http://www.sans.org>.
- [10] Bernhard Amann, Robin Sommer, Aashish Sharma, and Seth Hall. A lone wolf no more: Supporting network intrusion detection with real-time intelligence. In *15th International Conference on Research in Attacks, Intrusions, and Defenses*, pages 314–333, Berlin, Heidelberg, 2012. Springer-Verlag.
- [11] Eric M. Hutchins, Michael J. Cloppert, and Rohan M. Amin. Intelligence-driven computer network defense informed by analysis of adversary campaigns and intrusion kill chains. In L. Armistad, editor, *International Conference on Information Warfare and Security*, volume 6, pages 113–125. Academic Conferences International, Academic Publishing International Unlimited, 2011.
- [12] US Department of Defense. Joint Tactics, Techniques, and Procedures for Joint Intelligence Preparation of the Operational Environment (JP 2-01.3), June 2009.
- [13] Sergio Caltagirone and Deborah Frincke. ADAM: Active defense algorithm and model. In N.R. Wyler and G. Byrne, editors, *Aggressive Network Self-Defense*, pages 287–311. Syngress Publishing, Rockville, MD, USA, 2005.
- [14] Sergio Caltagirone. Evolving active defense strategies. Technical Report CSDS-DFTR-05-07, University of Idaho, Moscow, ID, USA, 2005.
- [15] Bruce Schneier. Attack trees. *Dr. Dobbs Journal*, 24 (12):21–29, 1999.
- [16] Richard Paul Lippmann and Kyle William Ingols. An annotated review of past papers on attack graphs. Technical Report PR-IA-1, Massachusetts Institute of Technology, Lexington Lincoln Laboratory, 2005.
- [17] Xinming Ou, Wayne Boyer, and Miles McQueen. A scalable approach to attack graph generation. In *Proceedings of the 13th ACM Conference on Computer and Communications Security*, pages 336–345. ACM, 2006.
- [18] Kyle Ingols, Richard Lippmann, and Keith Piwowarski. Practical attack graph generation for network defense. In *22nd Annual Computer Security Applications Conference, ACSAC'06*, pages 121–130. IEEE, 2006.
- [19] Lingyu Wang, Tania Islam, Tao Long, Anoop Singhal, and Sushil Jajodia. An attack graph-based probabilistic security measure. In *Data and Applications Security XXII*, pages 283–296. Springer, Berlin Heidelberg, 2008.
- [20] John Homer, Ashok Varikuti, Xinming Ou, and Miles McQueen. Improving attack graph visualization through data reduction and attack grouping. In *Visualization for Computer Security*, pages 68–79. Springer, Berlin Heidelberg, 2008.
- [21] Sebastian Roschke, Feng Gheng, and Christopher Meinel. Using vulnerability information and attack graphs for intrusion detection. In *Proceedings of the 6th International Conference on Information Assurance and Security (IAS 2010)*, pages 104 – 109, Atlanta, GA, USA, 2010. IEEE Press.
- [22] Vocabulary for event recording and incident sharing (VERIS). [ONLINE] <http://www.veriscommunity.net>.
- [23] ThreatConnect. [ONLINE] <http://www.threatconnect.com>.

- [24] A structured language for cyber threat intelligence information (STIX) . [ONLINE] <http://stix.mitre.org>.
- [25] John D. Howard and Pascal Meunier. Using a common language for computer security incident information. In Seymour Bosworth and M.E. Kabay, editors, Computer Security Handbook, chapter 3, pages 3.1–3.22. John Wiley & Sons, New York, NY, USA, 4th edition, 2002.
- [26] Frederick B. Cohen. Protection and Security on the Information Superhighway. John Wiley & Sons, New York, NY, USA, 1995.
- [27] Steven T. Eckmann, Giovanni Vigna, and Richard A. Kemmerer. STATL: An attack language for state-based intrusion detection. Journal of Computer Security, 10 (1) :71– 163, 2002.
- [28] Frederick B. Cohen. Information system attacks: A preliminary classification scheme. Computers and Security, 16 (1) :29–46, 1997.
- [29] John D. Howard and Thomas A. Longstaff. A common lanaguage for computer security incidents. Technical Report SAND98-8667, Sandia National Laboratories, October 1998.
- [30] Frederic Cuppens and Rodolphe Ortalo. LAMBDA: A language to model a database for detection of attacks. In Proceedings of the Third International Workshop, RAID 2000, pages 197–216, Berlin, Heidelberg, 2000. Springer-Verlag.
- [31] Michel Cedric and Ludovic Me. ADELE: An attack description language for knowledgebased intrusion detection. In Trusted Information, number 65, pages 353–368. Springer US, 2002.
- [32] Sophisticated indicators for the modern threat landscape: An introduction to OpenIOC. [ONLINE] [http://openioc.org/resources/An Introduction to OpenIOC. pdf](http://openioc.org/resources/An%20Introduction%20to%20OpenIOC.pdf).
- [33] Command and control. In Joint Publication 1-02: Department of Defense Dictionary of Military and Associated Terms, page 103. US Department of Defense, March 2009.
- [34] Wim Van Eck. Electromagnetic radiation from video display units: An evesdropping risk? Computers & Security, 4 (4) :269–286, 1985.
- [35] MITRE. Common vulnerabilities and exposures. [ONLINE] <http://cve.mitre.org/>.
- [36] Stuart McClure, Joel Scambray, and George Kurtz. Hacking Exposed. McGraw-Hill Osborne Media, 4th edition, 2003.
- [37] classtype. In Snort Users Manual 2.9.3, page 179. The Snort Project, May 2012.
- [38] Austin Troya, J. Morgan Groveeb, and Jarlath O’Neil-Dunne. The relationship between tree canopy and crime rates across an urban-rural gradient in the greater Baltimore region. Landscape and Urban Planning, 106 (3) :262–270, June 2012.
- [39] Will Gragido. Lions at the watering hole – the “VOHO” affair. [ONLINE] <http://blogs.rsa.com/lions-at-the-watering-hole-the-voho-affair>, July 2012.
- [40] urt Baumgartner. Winnti-stolen digital certificates re-used in current watering hole attacks on Tibetan and Uyghur groups. [ONLINE] [http://www.securelist.com/ en/blog/308194218/Winnti Stolen Digital Certificates Re Used in](http://www.securelist.com/en/blog/308194218/Winnti%20Stolen%20Digital%20Certificates%20Re%20Used%20in)

Current Watering Hole Attacks on Tibetan and Uyghur Groups, April 2013.

- [41] Matthias Vallentin, Jon Whiteaker, and Yahel Ben-David. The gh0st in the shell: Network security in the Himalayas. [ONLINE] <http://www.eecs.berkeley.edu/~yahel/papers/network-security-in-the-himalayas-cs294-28.pdf>.
- [42] Symantec Security Response. W32.Duqu: The precursor to the next Stuxnet. [ONLINE] [http://www.symantec.com/content/en/us/enterprise/media/security_response/whitepapers/w32 duqu the precursor to the next stuxnet. pdf](http://www.symantec.com/content/en/us/enterprise/media/security_response/whitepapers/w32_duqu_the_precursor_to_the_next_stuxnet.pdf), November 2011.
- [43] Red October: Diplomatic cyber attacks investigation. [ONLINE] [http://www.securelist.com/en/analysis/204792262/Red October Cyber Attacks Investigation](http://www.securelist.com/en/analysis/204792262/Red_October_Cyber_Attacks_Investigation), 2013.
- [44] Command Five Pty Ltd. SK Hack by an advanced persistent threat. [ONLINE] [http://www.commandfive.com/papers/C5 APT SKHack.pdf](http://www.commandfive.com/papers/C5_APT_SKHack.pdf), September 2011.
- [45] D. Ian Hopper and Richard Stenger. Large-scale phone invasion goes unnoticed by all but FBI. CNN, December 1999. [ONLINE] [http://edition.cnn.com/1999/TECH/ computing/12/14/phone.hacking/](http://edition.cnn.com/1999/TECH/computing/12/14/phone.hacking/).
- [46] Nate Anderson. How one man tracked down Anonymous – and paid a heavy price. Ars Technica, February 2011. [ONLINE] [http://www.arstechnica.com/tech-policy/ 2011/02/how-one-security-firm-tracked-anonymousand-paid-a-heavy-price/](http://www.arstechnica.com/tech-policy/2011/02/how-one-security-firm-tracked-anonymousand-paid-a-heavy-price/).
- [47] Jose Nazario. Georgia DDoS attacks – a quick summary of observations. [ONLINE] [http://ddos.arbornetworks.com/2008/08/ georgia-ddos-attacks-a-quick-summary-of-observations](http://ddos.arbornetworks.com/2008/08/georgia-ddos-attacks-a-quick-summary-of-observations), August 2008.
- [48] Brian Krebs. Cyber attacks target pro-Tibet groups. Washington Post, March 2008. [ONLINE] [http://www.washingtonpost.com/wp-dyn/content/article/ 2008/03/21/AR2008032102605.html](http://www.washingtonpost.com/wp-dyn/content/article/2008/03/21/AR2008032102605.html).
- [49] Bojan Zdrnja, Nevil Brownlee, and Duane Wessels. Passive monitoring of DNS anomalies. In Proceedings of the 4th International Conference on Detection of Intrusions and Malware and Vulnerability Assessments, DIMVA '07, pages 129–139, Berlin, Heidelberg, 2007. Springer-Verlag.
- [50] Manos Antonakakis, Roberto Perdisci, Wenke Lee, Nikolas Vasiloglou, II, and David Dagon. Detecting malware domains at the upper DNS hierarchy. In Proceedings of the 20th USENIX Conference on Security, SEC'11, pages 27–27, Berkeley, CA, USA, 2011. USENIX Association.
- [51] Wolfgang John and Tomas Olovsson. Detection of malicious traffic on back-bone links via packet header analysis. Campus-Wide Information Systems, (25) :342–358, 2008.
- [52] S. Templeton and K. Levitt. A requires/provides model for computer attacks. In Proceedings of the 2000 Workshop on New Security Paradigms, New York, NY, USA, 2001. ACM Press.
- [53] Crime pattern analysis: An investigative tool. In Michael J Palmiotto, editor, Critical Issues in Criminal Investigation, pages 59–69. Pilgrimage, 2nd edition, 1988.
- [54] Douglas M. Hawkins. The problem of overfitting. Journal of Chemical Information and Computer Sciences, (10) :1–12, 2004.
- [55] Huan Liu and Hiroshi Motoda. Feature Selection for Knowledge Discovery and Data Mining. Kluwer Academic Publishers, Norwell, MA, USA, 1998.

- [56] Susan W. Brenner. Organized cybercrime? how cyberspace may affect the structure of criminal relationships. North Carolina Journal of Law & Technology, 4 (1) , Fall 2002.