ある。表 3-10 に示すように、脅威のレベルと意図が最も高く、意図を実行する能力が最も高いものを「重要」とする。そして、全体的な脅威は「深刻」へとスケールダウンしていく。最も懸念の少ない脅威は、「無視できる（Negigible）」と分類される。

表 3-10．脅威レベルマトリックス脅威レベルマトリックス．

意図のレベル	能力レベル					
		非常に小さい	小規模	限定的	重大な	手強い
	非常に高い	中程度	中程度	シビア	シビア	クリティカル
	高	低	中程度	大幅な	シビア	クリティカル
	中位	低	低	中程度	大幅な	シビア
	低	無視できる	無視できる	低	中程度	大幅な
	非常に低い	無視できる	無視できる	低	低	中程度

この可能性スコアを、ある脅威者が攻撃を成功させた場合に起こりうる結果と照らし合わせて、リスクの高さを判断し、他の特定されたリスクに対して優先的に対策を講じるべきかどうかを決定することができる。

(6) 偽装工作の可能性の判断

欺瞞、または難読化とは、悪意のある脅威行為者がその身元、目標、技術を隠すために使用する TTP のことを指す。防御者が活動を特定するための手がかりを残さないようにするため、脅威者はインターネット上でデータを密かに送信するツールやテクニックを使用することができる。

また、洗練された脅威行為者は、偽旗作戦を行うことがある。これは、ある行為者が他の行為者の既知の活動を模倣することで、防御者にその活動を他の行為者のものと誤認させることを狙ったものだ。例えば、ある国家は、サイバー犯罪者や他の国家が広く使用していると思われるツールを使用し、攻撃が無関係の行為に起因すると誤認されることを期待することができる。

サイバー犯罪者が自らの行動をうまく隠蔽する能力は、その巧妙さと動機づけのレベルに応じて異なる。一般に、国家や有能なサイバー犯罪者は、他の脅威行為者よりも難読化に長けている。

最近の悪意のあるサイバー行為者の手法には、「欺瞞」が含まれる。悪質なサイバー行為者は、偽装した別の ID を採用することで活動を隠したり、別の国にサイバー攻撃の濡れ衣を着せたりすることもある。一連のスパイ行為では、他国のハッキングインフラを乗っ取り、被害者をスパイしたり、マルウェアを配信するために利用したことがある。例えば、ロシアのハッカー集団「Turla」や「Waterbug」は、「OilRig」と呼ばれるイランのハッカー集団のサーバーを乗っ取り、ロシアの目的に利用するという複雑

なスパイ行為を長年にわたって行っている。[8]

3. アトリビューションプロセスを支援するモデル

CTI フレームワークの主な利点は、組織が敵の活動方法、および敵が最初のアクセスの獲得、発見、横方向への移動、データの流出を計画する手順について理解できるようになることだ。これにより、攻撃者の視点から活動を見ることができ、アトリビューション理論を構築する際に不可欠な動機と戦術をより深く理解することができる。さらに、組織はこの理解と知識を活用して、セキュリティ体制のギャップを特定し、脅威の検知と対応を改善することができる。これは、チームが攻撃者の次の行動を予測し、迅速に対処することを可能にする。

さらに、サイバーセキュリティのスキル不足が深刻化している現在の職場環境において、このフレームワークは、若手や新規採用のセキュリティスタッフに必要な知識と調査ツールを提供し、脅威データベースの構築をサポートする組織内のすべてのセキュリティ専門家の集合知を活用して、特定の悪質な脅威要因に迅速に対応できるようにすることが可能だ。

(1) ダイアモンド・モデル

セキュリティ・チームは、ネットワーク侵入の「誰が、何を、いつ、どこで、なぜ、どのように」するかを理解し、進行中の攻撃への対応と、攻撃を事前に軽減するアプローチの両方を開発する必要がある。これらの質問に対する回答の価値を高めるために、ネットワーク防御者は、脅威のデータを合成し、関連付け、文書化する能力が必要だ。

侵入分析のダイヤモンド・モデルは、これを実現するための 1 つのフレームワークだ。このアプローチでは、すべてのインシデントをダイヤモンド型のレンズを通して見る。ダイヤモンドの 4 つの要素（敵対者、能力、インフラ、被害者）は、攻撃の関係性と特徴を特定し、強調する。これら 4 つの中核的要素を検証することで、特定の悪意ある行為に関する洞察を得て、知識を得ることができる。4 つの要素を以下に示す。

- 敵対者：目標を達成するために、被害者に対してある能力を活用する責任を負う組織または脅威行為者である。
- ケイパビリティ：敵対者があるイベントで使用するツールやテクニックを指す。
- インフラストラクチャ：敵対者がケイパビリティを提供するために用いる、インターネット・プロトコル（IP）アドレスや電子メールアドレス、ドメイン名などの物理的または論理的な通信構造を含む。
- 被害者：攻撃が開始され、脆弱性が悪用され、または能力が使用される標的のこと。被害者は、組織、人、またはターゲットの電子メールや IP アドレス、ドメインなどの資産である。

要約すると、侵入分析のダイヤモンド・モデルは、「敵対者」が「被害者」に対して「インフラ」上で「能

[8] Turla Espionage Group Hacks OilRig APT Infrastructure, 2019, https://www.bleepingcomputer.com/news/security/turla-espionage-group-hacks-oilrig-apt-infrastructure/

力」を使用することを説明している。このモデルの原理は、以下の図 3-8 に示すように、すべての侵入に対して、敵対者は被害者に対してインフラ上の能力を活用し、インパクトを与えることで目標に向かって進むというものだ。

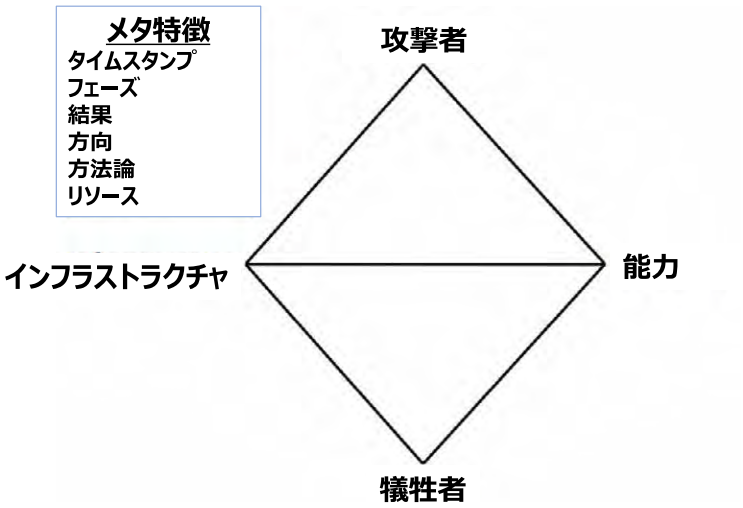

図 3-8：分析のダイアモンド・モデル.[9]

ダイヤモンド・モデルは、文脈に応じた指標を提供することにより、脅威情報の共有と他の計画フレームワークとの統合を向上させる。また、インテリジェンスのギャップを検出し、サイバー分類法、オントロジー、脅威インテリジェンスの共有プロトコル、および知識管理の基礎を築く。さらに、仮説の生成、テスト、文書化を促進することで、セキュリティ・チームは分析の精度を高め、分析プロセスの精度を向上させることができる。

ダイヤモンド・モデルの使用例

- 分析的ピボッティング

ピボットとは、あるデータ要素を取得し、データソースと連携してそれを活用し、他の関連する要素を特定する分析手法である。ピボットとは、仮説検証のことである。ピボッティングの成功は、セキュリティアナリストが要素間の関係を理解し、データ要素とそのソースを活用する能力を備えているかどうかにかかっている。

- 知識ギャップの発見

イベントに含まれていないダイヤモンド・ノードや、アクティビティ・スレッドに欠落しているイベントを、ダイヤモンド・モデルで接続することができる。これにより、ナレッジギャップを特定し、インシデン

[9] Sergio Caltagirone, Andrew Pendergast, and Christopher Betz, "Diamond Model of Intrusion Analysis," Center for Cyber Threat Intelligence and Threat Research, Technical Report ADA586960, 2013.

ト対応と脅威のインフラおよび能力に焦点を当てることができる。

- 中心型アプローチ (Centered Approach

中心型アプローチは、ダイヤモンド・モデルの特定の機能に焦点を当て、新たな悪意のある活動を検出し、他の関連する機能に影響する活動を公開する。中心型アプローチには、敵対者中心型、能力中心型、インフラ中心型、被害者中心型、社会・政治中心型、技術中心型という 6 つの中心型アプローチがある。最初の 4 つはダイヤモンドのノードに焦点を当て、残りの 2 つはダイヤモンドのメタフィーチャーに焦点を当てる。

ダイヤモンド・モデルはどのような場合に有効か？

- 異なる侵入を比較し、グループ化する
- 一見、異質な活動間の類似性を検証する

ダイヤモンド・モデルの限界

- 高次のレベル
- 柔軟性が高すぎる - 情報をどのように「ビン詰め」するかは、ユーザー自身がチーム内で決める必要がある

ダイヤモンド・モデルはアトリビューションを判断する上で非常に重要なモデルである。ダイヤモンド・モデルは意思決定者やリーダーとの帰属に関する議論の枠組みを作るのに役立つ。

(2) Q モデル

Thomas Rid と Ben Buchanan の論文「Attributing Cyber Attacks」で紹介された Q モデル (図 3-9) は、アトリビューションが一本の直線的な経路ではないことを強調したものである。その目的は、技術的な詳細とアトリビューションに使用される方法を提供することで、アナリストがより多くの情報に基づいた疑問を持ち、結論に疑問を持てるようにすることである。

Qモデルは4つの情報レベルを持っている。黒字の外側のレベルは戦術的、灰色の中間のレベルは作戦的、白字の内側のレベルは戦略的である。最後のレベルは、Qの「フック」であるコミュニケーションである。例えば、アナリストがする質問は、レベル別に分けることができる。「何を」「どのように」は戦術的、「誰が」は作戦的、そして「なぜ」は戦略的となる。そして、これらの質問に対する答えは、攻撃への対応に関する意思決定を行うために、リーダーシップやその他の利害関係者に伝達するために使用される。

Qモデルは、アトリビューションの質が、アナリストの質問、方法、プロセス全体の概観の関数であることを強調する。Qモデルは、これらの構成要素を調査プロセスの4つのレベルで考えるための有用な構造を提供する。

- 戦術的/技術的レベル：何が／どのようにして攻撃が起こったのか、技術的な疑問を提起する。侵害の指標、侵入経路、ペイロード、ネットワーク活動など、技術的な証拠を評価する。
- 運用レベル：情報の統合により、何が起こったのか、より高度なアーキテクチャを理解し、攻撃の犯人を突き止める。これには、攻撃の技術的精巧さの評価、国家および非国家主体による既知の能力との比較、事件の地政学的背景の理解などが含まれる。

- 戦略的レベル：なぜそのような攻撃が行われたのかを判断する。このレベルでは、結論を導き出すためのストレステストを行い、悪意のある出来事の根拠を理解しようとし、一連の出来事が意味のある前例となったかどうかを判断する必要がある。
- コミュニケーション・レベル：帰属の結論がどのように伝達されるべきかを記述している。このモデルでは、より詳細な情報、推定的な表現、分析の限界などを伝えることで、より優れた集団防衛を可能にし、帰属の信頼性を高め、帰属自体を向上させることを主張している。

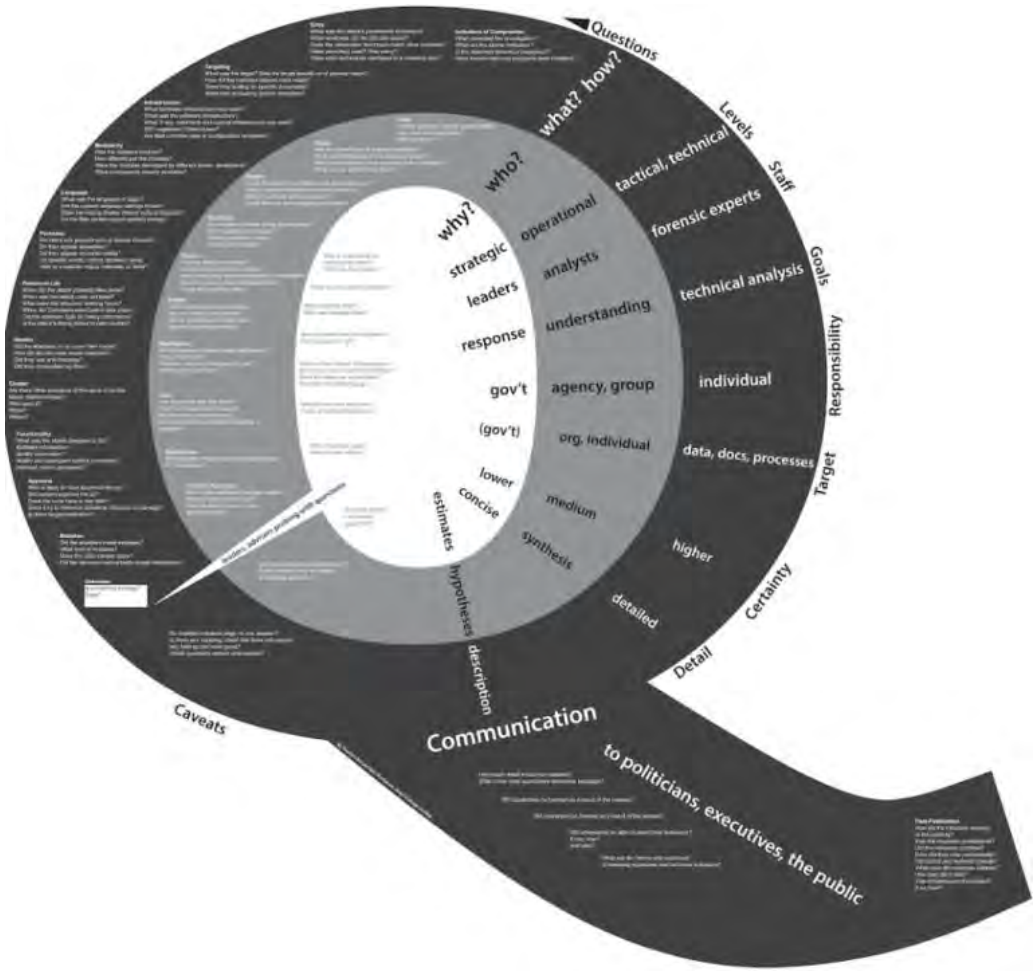

図 3-9. Attributing Cyber Attacks における Q モデル.[10]

[10] Thomas Rid and Ben Buchanan, "Attributing Cyber Attacks", Journal of Strategic Studies, Volume 38, 2015

Q Modelの限界

- Qモデルは、多くの組織や民間企業で採用されていないため、運用に制限がある。
- 国家安全保障上の脅威となる事象のアトリビューションに焦点を当てすぎており、犯罪的なサイバー侵入の事象のアトリビューションには焦点が当てられていない。
- 国家によるアトリビューションを公にすることでメリットが得られる場合もあるものの、敵対者を抑止するために公的なアトリビューションが必要なわけではない。

(3) MITRE ATT&CK モデル

MITRE ATT&CK (図3-10) は、実世界の観察に基づく敵対者の戦術と技術に関するグローバルにアクセス可能な知識ベースである。 MITRE ATT&CKは、APTグループがサイバー攻撃で使用した悪質な行為に関する情報をまとめている。ATT&CKは、Adversarial Tactics, Techniques, and Common Knowledgeの略で、これらのグループのTTPの詳細な説明が含まれている。ATT&CKの知識ベースは、民間企業、政府機関、サイバーセキュリティ製品・サービスコミュニティにおいて、特定の脅威モデルや方法論を開発するための基盤として使用されている。MITREは、ATT&CKが敵対者の行動の詳細な情報を提供し、特定の攻撃の背後にいる人物に関する結論をサポートするために使用される事例を文書化している。

MITRE ATT&CKは、敵対者が使用するTTPの説明、特定の技術に対する検出のための提案や一般的な緩和策の提供、APTグループの既知の手法、特徴、特定の攻撃起因のプロファイリングを行う百科事典のようなものだ。ATT&CKはまた、攻撃に使用されるソフトウェア（マルウェア、合法的または悪意を持って使用できる市販およびオープンソースのコードの両方）の広範なリストも提供する。ATT&CKに取り込まれたすべての情報は、一般に公開されているデータやレポート、およびコミュニティから提供されたものだ。最新版のATT&CK for Enterpriseには、14の戦術、188のテクニック、379のサブテクニック、129のグループ、および637のソフトウェアに関連する情報が含まれている。

ATT&CKは、世界中の知識豊富なサイバー専門家によって提供され、攻撃者のタイプやカテゴリー別に構成されているため、日本政府に最も関連するものを含め、攻撃者が使用しているTTPを特定するための重要なリソースとなっている。

ATT&CKは、世界中の洗練されたネットワーク防御者にとって、脅威に関する情報を確実に入手するためのCTIパズルの基礎となる部分となりつつある。例えば、英国のサイバーセキュリティ情報共有パートナーシップ（CiSP）は、ATT&CKの知識ベースを情報共有プログラムに統合する最善の方法について取り組んでいる。日本を含む多くの民間企業は、ネットワーク防御を強化するために、TTPに関するATT&CKの情報を活用することの重要性を認識している。

ATT&CKはどんな時に役に立つのか？

- 敵対者の行動を詳細なレベルで追跡する
- 特定の行動に関して、防衛側や他の組織と共通言語でコミュニケーションをとることができる

ATT&CKの制限事項とは？

- CTIの全側面や全技法をカバーしているわけではない
- 戦術的な側面が強い
- 複雑で、学習曲線が急な場合がある

MITRE は、ほとんどのモデルやフレームワークが ATT&CK を補完するものであり、ネットワーク防御者が 1 つのフレームワークだけを選択する必要はないと考えている。例えば、ATT&CK とダイヤモンド・モデルは補完関係にある。ATT&CK は敵の詳細な行動を記録し、Diamond Model は防御側が侵入のクラスタリングを行う場合に役に立つ。また、両者を併用するケースもある。例えば、ATT&CKでマップされた技術は、敵の能力を分析するためのダイヤモンド・モデルへの有効な入力ソースとなり得る。MITRE ATT&CK フレームワークは有用であるものの，それだけではアトリビューションの結論を支持するのに十分ではないこともある。MITRE ATT&CKは、ダイヤモンド・モデルなどの他の分析モデルと組み合わせることで、アトリビューションに関する結論をより確かなものにすることができる。

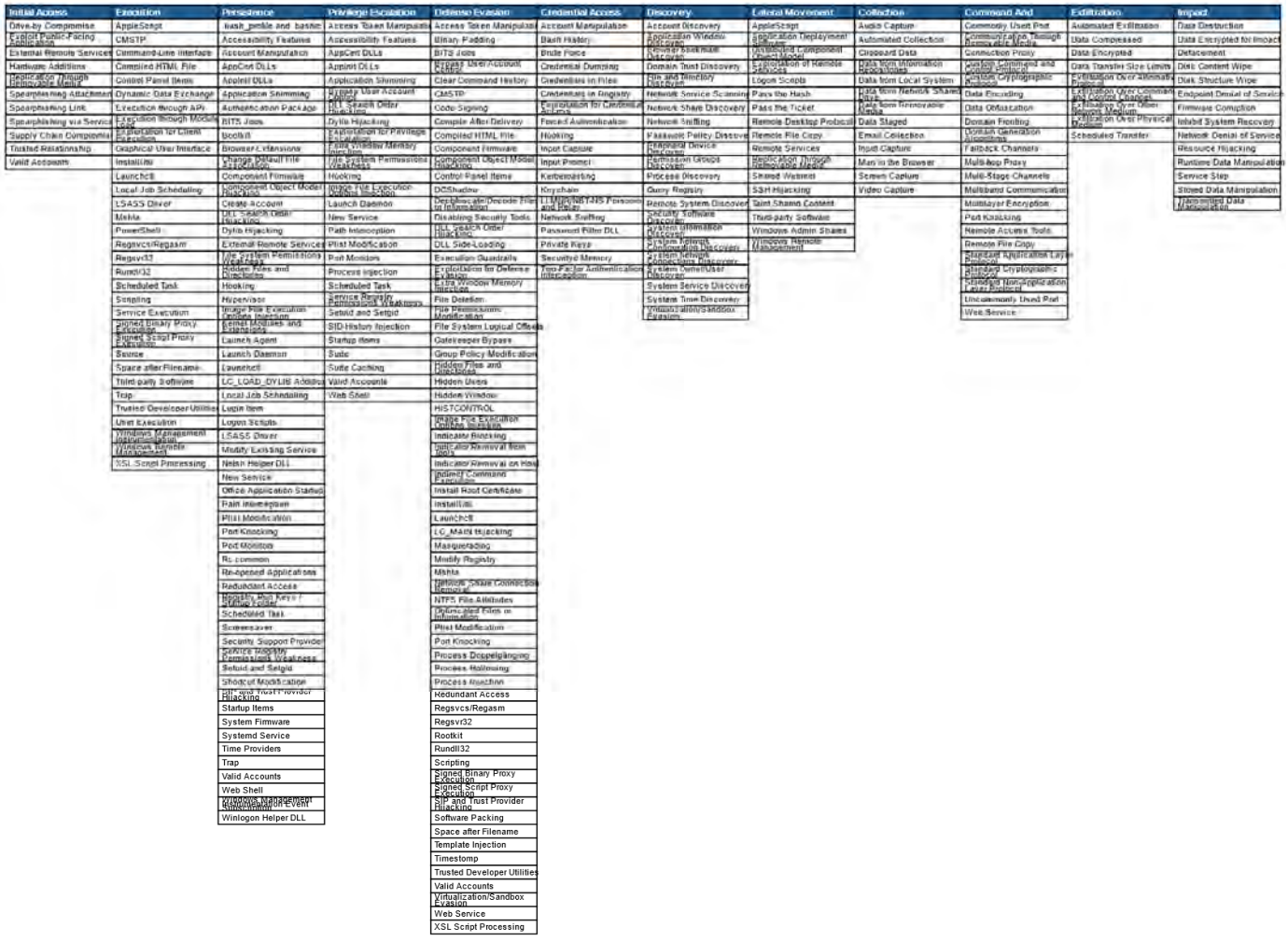

図 3-10. MITRE ATT&CK フレームワークの例.

4. アトリビューションのための分析ツール

前述のとおり、悪意のある活動に対する責任について正当な結論を導き出すには、一連のツール、技術、および分析が必要だ。ネットワーク上で観測されたアクティビティの責任を確認しようとする個人にとって、次のような種類の分析が有用である。

(1) マルウェア分析

マルウェアは、あらゆる種類の悪意のある不要なソフトウェアを指す一般的な用語だ。マルウェアは、サイバースペースにおける最も主要な犯罪の原因となっている。悪意のあるコードは、物理的なアクセスまたはネットワーク手段によってホストに侵入することができる。マルウェアベースの解析による悪意のある行為者の特定は、実際の攻撃者の特定が複雑であるため、依然として困難だ。ほとんどの場合、マルウェアベースの解析は、使用された悪意のあるソフトウェアまたはコードを特定するために利用される。また、既知の攻撃との類似性が確認された場合、マルウェアベースの解析により攻撃者の地理的な位置の特定に成功する場合もある。

(2) 静的解析

マルウェアの解析には、静的解析と動的解析がある。静的解析では、コードは実行されずに解析される。プログラムのソースコード、または逆アセンブルされたバイナリのいずれかになる。その後、悪意のあるコードを検出するために、コードを既知のシグネチャと比較する。静的解析は、難読化技術のために制約があり、解析されたコードが実際に実行されるコードでない可能性がある。

(3) 動的解析

動的解析では、仮想マシン（VM）のような制御された環境でコードを解析する。VMを使用することで、解析のために悪意のあるコードを実行すること自体が重大な懸念となるため、安全性が確保さる。動的解析では、悪意のあるコードをVM環境で直接実行できるため、静的解析手法の多くの制限を克服することができる。

(4) 類似性アトリビューション

類似性ベースのアトリビューションにより、悪意のあるイベントで使用されたマルウェアは、以前の攻撃で使用されたマルウェアと比較することができる。例えば、Kaspersky Labのリサーチャーである Kurt Baumgartner氏は、2014年のソニーへの侵入と、一般的に北朝鮮が関与したとされる他の事件との間に、いくつかの類似点があることを指摘する。Baumgartner 氏は、攻撃者は、「Destover」と呼ばれる破壊的なワイパー型マルウェアを展開し、全社的にハードディスクの上書きを行うことで痕跡を消したことを指摘する。また、「同じマルウェアが、韓国を標的としたDarkSeoul攻撃で使用されたと報告されており、これは同国の北の隣国によるものとされている」とも述べている。類似性ベースのアトリビューションの大きな限界は、以前に知られていたマルウェアとの類似性が、必ずしも以前にアトリビューションされた

攻撃者の攻撃であるとは限らないということだ。これは、以前に使用されたマルウェアのシグネチャを盗用またはコピーして、攻撃のアトリビューションを誤らせることも可能なためだ。さらに、多くの悪意ある攻撃者は、市販のソフトウェア（COTS）を利用することが知られており、その起源をさらに難解なものにしている。

(5) 間接的なアトリビューション方法

敵対者は、マルウェアのコードを難読化したり、偽造または盗難されたIDを使用したりできるため、ネットワーク・トレースバックなどの直接的なアトリビューション技術には限界がある。代替手法として、間接的なアトリビューション技術を検討する必要がある。特に、複数のコンピュータを標的とした多段攻撃や犯罪に当てはまる。間接的帰属とは、攻撃者の行動に関する統計的モデルを使用して、悪意のあるサイバー事象（または犯罪）を攻撃者（または犯罪者）に帰属させるプロセスだ。これらの行動モデルは、文体、ソーシャルネットワーク分析、コーディングの類似性など、さまざまな属性に基づいて構築される。これらの特性は互いに関連付けられ、悪意のある行為者のプロファイルを生成するために使用できる。間接的アトリビューションでは、ニューラルネットワーク、遺伝的アルゴリズム、サポートベクターマシンなどの技術を組み込んで、犯罪者プロファイルを生成できる可能性がある。しかし、脅威となる行為者の正確なプロファイルを生成するためには、広範なデータを利用できることが必要だ。

(6) 機械学習技術

機械学習技術は、サイバー犯罪のアトリビューション問題を解決するために、悪意のあるソースを特定するために使用することができる。このような技術は、長期間にわたってネットワークログを収集し、それを分析して、悪意のある活動に関与しているIPソースの集合を特定することに依存している。分析は、クラスタリング、ニューラルネットワーク、サポートベクターマシンなど、さまざまな機械学習技術によって行うことができる。

(7) 行動分析

Dacierらは、クラスタリング技術を適用して、信頼できるIPと悪意のあるIPのソースを特定した。この情報は、サイバー犯罪の発生源を特定するために適用された。行動分析は、悪意のあるソースを特定する上で効率的だが、膨大な量のデータを利用できることに依存する。そのため、他の手法と組み合わせて使用されることも少なくない。例えば、IPアドレスクラスタリングの場合、ハニーポットやウェブサーバーなど、さまざまなソースからログを収集することが可能だ。このような手法の大きな制約として、スプーフィングや匿名化手法の存在により、マルウェアの発生源の信憑性が低くなってしまうことが挙げられる。

(8) 遺伝的アルゴリズム

遺伝的アルゴリズムも、マルウェアのアトリビューションに使用されている。この手法では、マルウェアの遺伝学、進化の過程、およびその特性を利用して、攻撃者の出自を特定し、将来の攻撃の特性を予測する。さらに、攻撃の意図も考慮し、包括的なシステムを構築している。行動解析や機能解析を用いるこ

とで、マルウェアの出所や機能、系統の特徴など、いくつかの機能を特定することができる。しかし、遺伝子解析は、マルウェアの特徴を把握する必要があるため、簡単にはいかない。

(9) ニューラルネットワーク

ニューラルネットワークに基づく手法も、作者の特定に使用されている。この手法には、電子メールなどの文書の実際の作成者を特定することも含まれる。電子メールのアトリビューションは、スパム、フィッシング、サイバーテロに関連するサイバー犯罪に必要とされる。サポートベクターマシンやニューラルネットワークなどの技術を用いることで、異なる文体を識別してクラスタリングし、電子メールの作成者を特定することができる。つまり、電子メールのような文書が特定の人物に由来すると断定することはできない。さらに、電子メールは複製される可能性があり、アトリビューションを決定的なものにすることはできない。

(10)ソーシャルネットワーク

ソーシャルネットワークの準識別子（QID）を利用することによってもアトリビューションを試みることができる。QID（性別や郵便番号など）は、公共データセットから何らかの情報を明らかにするための識別子である。QIDは一意な識別子ではないが、他のQIDと組み合わせることで一意な識別子を作成することができる。ソーシャルネットワークは大規模なネットワークで構成されているため、ソーシャルネットワークからの膨大なデータセットを活用し、サイバー攻撃を悪意のあるサイバー行為者に正確にアトリビューションするためには、膨大なデータと分析が必要となる。さらに、この手法の重大な懸念は、ハッカーが偽の識別子を使用して不正な目的を達成することができることだ。

間接的アトリビューション技術を強調する学術研究論文は数多くあるが、それらはシミュレーションされた事象に依存しており、現実の事象に効果的に適用されたことはない。また、これらの技術を有効に活用するためには、大量のデータが必要であり、現在までのところ、この量は容易に入手できないことに注意する必要がある。マルウェアや悪意のあるサイバーイベントの検出、特定、アトリビューションを支援するための機械学習（ML）や人工知能（AI）の活用など、高度な技術を論じた研究論文があるものの、これらの研究のほとんどは、限られたデータと現実のサイバー運用に熟練していない人材がいるアカデミックな環境で行われていることが問題である。ML、AI、ニューラルネットワークの使用に関する研究は今後も継続され、さらにデータが利用可能になれば、これらの技術が悪意のあるサイバーイベントの検出、防止、帰属の強化を支援することが期待される。

以下の表3-11は、これらの手法の概要、手法の簡単な説明、および各手法の可能な制限を示したものである。