L-021

実行可能な攻撃プランに基づくリスクアセスメント
Risk-assessment based-on feasible attack plans

遊座 広太郎† 八槇 博史†
Kotaro Yuza Hirofumi Yamaki

1. はじめに

近年、リスクアセスメントは様々な分野において重要な役割を果たしており、その手法は多岐にわたる。その中でリスクアセスメント手法の一つとしてチェックリスト利用法がある。チェックリストはガイドラインに基づいて作成されるが、IT システムのリスクアセスメントにおいては、現実的でない攻撃が過大評価されることがある。

本論文では、リスクアセスメントの全体像を概観し、既存の手法とその分類について紹介する。そして、チェックリスト利用法の課題を指摘して解決する手法を提案する。これらを通じて、リスクアセスメントの現状を理解し、本論文で提案する手法の必要性とその位置づけを明確にする。

提案手法は、MITRE ATT&CK[1]などの要素を取り入れたプランニングによるリスクアセスメント手法である。この手法では、攻撃者が目的を達成するための Techniques の系列を自動的・論理的に出力し、実現可能性のある攻撃のみを評価対象とすることを可能とする。

2. 現在のリスクアセスメント

2.1 概要

ISO31000:2018(JIS Q 31000:2019)[2]において、リスク(risk)は「目標に対する不確かさの影響」とされ、リスクアセスメント(risk assessment)は「リスク特定、リスク分析及びリスク評価を網羅するプロセス全体」として定義されている。

リスク特定は、リスク源となる脅威、影響を受ける資産もしくは資産グループ、脅威が発生した際に生じる事象とその結果を特定するプロセスである。

リスク分析は、リスク特定において特定したリスクについて、定性もしくは定量的な分析を行い、リスクレベルを決定するプロセスである。

リスク評価は、リスク分析の成果に基づき、どのリスクへの対応が必要か、対応の優先順位をどうするかについて、意思決定者に必要な情報を導き出すプロセスである。

以上のプロセスからなるリスクアセスメントによって、組織は潜在的な脅威や脆弱性を特定し、その影響を評価し適切な対策を講じることが可能である。

2.2 既存手法と分類

リスクアセスメント手法の分類は、GMITS に記載された表 1 に示す 4 分類が広く知られている。

ベースラインアプローチは、共通のリスク対策(ベースライン)をチェックリスト化して調査する方法である。チェックリスト利用法はその一つである。これは過去の知見などによって作成された確認項目の一覧を用いて、リスクアセスメントを行う方法である。佐々木ら[3]は、チェックリストがうまく作成されているときには、項目を理解できれば専門知識をもたない人でも利用することができること、チェック項目を共通化することによりリスクアセスメントの結果の比較や第 3 者によるレビューが可能なことを利点として挙げている。

詳細リスク分析アプローチは、資産ごとに関連するリスクを識別し、リスクを算定する方法である。多くの専門知識や労力が必要となるが、対象を詳細に分析し、個々に最適なリスク対策を実施したい場合に有効である。ランク値付きマトリックス法や故障木解析などがこれに含まれる。

組合せアプローチは、一般にベースラインアプローチと詳細リスク分析アプローチを組み合わせた方法である。上位レベル分析を行い、重要な対象領域に対して詳細リスク分析を行い、その他の領域に対してはベースラインアプローチを用いる。

非形式的アプローチは、熟練した専門家らの知識と経験を用いて、主観的にリスクを洗い出す方法である。

表 1　リスクアセスメント手法の分類

種類	概要
ベースライン アプローチ	既存の標準・基準や事例などから作成したチェック項目を利用
詳細リスク分析 アプローチ	対象となる資産ごとに、構造化され解析的な手法により分析
組合せ アプローチ	ベースラインと詳細リスク分析を組み合わせて実施
非形式的 アプローチ	専門家の知識や経験を基に主観的に分析

2.3 チェックリスト利用法の課題

佐々木らは、チェックリスト利用法の利点を述べると共にリスクアセスメントの結果がチェックリストのできに依存していること、想定外のリスクに対する考慮が難しいこと、リスクレベルの定量的もしくは定性的分析ができないことを欠点として挙げている。

本論文では「チェックリストのでき」に関して、「対象システムへの現実的でない攻撃が過大評価される」という点に注目する。

時間やコストなどリスク対策を講じるためのリソースは有限である。よって、現実的でない、すなわち実現可能性のない攻撃に対して対策を講じることは望ましくない。しかし、チェックリストでは実現可能性のない攻撃が過大評価され、リスク対策の決定に適さない場合がある。

具体的なチェックリストの一つとして IPA(情報処理推進機構)が公開している「5 分でできる！情報セキュリティ自社診断」[4]がある。これは同じく IPA が公開している「中

†東京電機大学大学院 システムデザイン工学研究科 情報システム工学専攻 (Information System Engineering, Graduate School of System Design and Technology, Tokyo Denki University)

小企業の情報セキュリティ対策ガイドライン第 3.1 版」[5]の中でもベースラインアプローチとして紹介されている。例えば、「新たな脅威や攻撃の手口を知り対策を社内共有する仕組みはできていますか？」というチェック項目では、「一つ一つのホストに関して新たな脅威や攻撃の手口を知り対策を社内共有する仕組みが無い場合に起こり得る攻撃」を過大評価している。リソースが限られた中では、攻撃者が目的を達成するまでに用いないホストについては対策を講じる必要はない。よって、攻撃者が用いるホストのみをチェックリストの対象とすることが望ましい。

このような点を解決し実現可能性のある攻撃のみを評価の対象とすれば、より適切なリスクアセスメントが可能である。

3. 提案手法

本セクションでは、指摘したチェックリスト利用法の課題を解決する提案手法の概要を説明する。次に具体的な実装方法を示し、提案手法の有効性を検証するためのシナリオとその結果及び評価について述べる。

3.1 概要

提案手法は、MITRE ATT&CK などの要素を取り入れたプランニングによるリスクアセスメント手法である。

MITRE ATT&CK は攻撃者の攻撃手法や行動、目的を明文化したフレームワークである。そして、プランニングは人工知能の一つであり、とりうるアクション群があるときに特定の初期状態から目標状態に到達するまでの一連のアクション(プラン)の導出を目的とする。今回は、STRIPS（Stanford Research Institute Problem Solver）を使用した。STRIPS では各アクションが、前提条件とそのアクションが実行された場合の効果として表される。

本手法では STRIPS で攻撃者が目標を達成するまでのアクションの系列をプランとして求めることで実現可能性のある攻撃のみを評価の対象とする。MITRE ATT&CK の情報をもとに攻撃手法の前提条件と効果を記述してアクションを定義する。このため、出力されたプランに関わるホスト、つまり攻撃の前提を満たすホストのみをチェック項目の適用対象とすることができる。

攻撃の前提を満たすホストかどうかを判断することは専門知識をもたない人には困難であり、この部分は非形式アプローチに該当すると考えられる。しかし、本手法ではプランニングの入力ファイルを記述さえすればプランナーが自動で判断しチェック項目を適用すべき箇所を自動的・論理的に出力する。よって、本手法をチェックリスト利用法と組み合わせて使うことで実現可能性のない攻撃の過大評価部分を削減することが期待できる。

3.2 実装

プランニングは基本的に最適なプランを 1 つだけ導出する。そのため、攻撃者が目的を達成するプランが複数ある場合に対応できない。そこで今回はアクションごとにコストを設けて合計コストが最小となるプランを導出した。導出されたプランに関わるアクションのコストを都度増加することで再び同じプランが導出されるまでプランニングを繰り返し、攻撃者が目的を達成するプランを全て導出した。

プランニング用のファイルを記述する言語である PDDL がアクションに対してコストを設定することをサポートしたのは PDDL3.1 からである。よって、今回は PDDL3.1 を使用できるプランナーとして Fast Downward[6]を使用した。

今回、プランニング用のファイル(以降 PDDL ファイルと呼称する)を Domain と Problem の 2 つのファイルに分ける。Domain ファイルでは、アクション(Action)や状態を表す述語(Predicate)を定義する。Problem ファイルではプランニングで解く対象となる問題、すなわち初期状態と目標状態を定義する。

Predicate は図 1 に示すように状態を表しアクションの前提条件と効果などに用いられる。MITRE ATT&CK では攻撃手法を Techniques として整理しており、それぞれに ID が振られている。例えば、「Exfiltration Over C2 Channel」という Technique は ID が「T1041」であるため、攻撃者によりこの Technique が達成された状態を「achieved_T1041」と定義した。また、T1041 が属する Tactics「Exfiltration」の ID は「TA0010」である。したがって、この Tactics に属する Technique が達成された状態は「achieved_TA0010」とした。

```
(:predicates
    (controllable ?x - host)
      ; ?x が完全に攻撃者の制御下にある
    (communication ?from - host ?to - host)
    ; ?from から ?to へ通信が許可されている
    (has_credential ?attk - host ?target - host)
      ; ?attk が ?target の credential を所持している
    ; (achieved_XXX ?x - host)
      ; ?x において XXX が達成されている
    (achieved_T1041 ?x - host)
    (achieved_TA0010 ?x - host)
)
```

図 1　Predicate の記述例

Action は図 2 に示すように攻撃手法を表している。前提条件(precondition)と効果(effect)を Predicate で表している。

```
(:action T1041 ; Exfiltration Over C2 Channel
    :parameters (?attk - host ?target - host)
    :precondition (and
      (controllable ?attk)
      (communication ?attk ?target)
      (has_credential ?attk ?target)
    )
    :effect (and
      (achieved_T1041 ?target)
      (achieved_TA0010 ?target)
      (increase (total-cost) 1)
    )
)
```

図 2　Action の記述例

例えば、T1041 は「Exfiltration Over C2 Channel」という Technique であり、「既存の C2 チャネルを介してデータを盗み出す」という内容である。これを PDDL ファイルに Action として記述すると図 2 のようになる。事前条件とし

て、ホスト?attk が攻撃者の制御下にあり、データを盗み出す攻撃対象となるホスト?target へ?attk から通信が可能かつ、?attk が?target に対する認証情報を持っていると定義した。Action が実行された場合の効果は、T1041 と TA0010 が達成された状態となりコストが 1 増加するとした。導出されたプランにこの Action が含まれた場合はコストの増分を大きくして再度プランニングを行い、導出済みのプランが出るまでコストの増加とプランニングを繰り返す。

Problem ファイルには、リスクアセスメントの対象となる被攻撃システムの状態を初期状態、攻撃者が目的を達成した状態を目標状態として Predicate で記述する。図 3 に例を示す。

```
(:objects
    hostA - host
    hostB - host
)
(:init
    (communication hostA hostB)
    (communication hostB hostA)
    (has_credential hostB hostA)
)
(:goal (and
    (achieved_TA0010 hostA)
    )
)
(:metric minimize (total-cost))
```

図 3　Problem ファイルの記述例抜粋

この例では、リスクアセスメントの対象となるシステムが 2 つのホスト hostA と hostB によって構成されており、相互に通信可能かつ hostB が hostA の認証情報を持っているという初期状態を与えている。目標状態は攻撃者が hostA に対して TA0010 を達成した状態としている。この場合、図 2 の T1041(Action)に従うと、(controllable hostB)を満たしていなければプランは生成されない。つまり hostB が攻撃者の制御下になければ攻撃者は T1041 を使用して目標を達成することはできないと論理的に決定される。

以上のように Domain と Problem の 2 つの PDDL ファイルを記述してプランニングを行い、実現可能性のある攻撃のみを評価してチェック項目を適用すべき箇所を自動的・論理的に出力する。

3.3 検証

提案手法の有効性を検証するために、具体的なシナリオを作成し、プランニングを行った。その内容と結果について述べる。

3.3.1 シナリオ

ある中小企業が図 4 に示すシステムを利用している場合を想定し、ファイルサーバーの情報流出を事前に防ぐためチェックリストの対象とすべきホストを導出する。リスクアセスメントの対象となるシステムは dmz と office_lan の 2 つのネットワークを持つ。

dmz では、外部からの通信が許可されている 2 つの WEB サーバー(public_web_srv1、public_web_srv2)と 1 つのメールサーバー(mail_srv)が置かれている。そして、リモートワークを行う従業員用の VPN サーバー(vpn_srv)も同じく dmz に置かれている。

office_lan ではファイルサーバー(file_srv)、Web サーバー(private_web_srv)、DNS サーバー(internal_dns_srv)が置かれており、オフィス内で働く従業員の PC(office_worker_pc)も office_lan 上に存在する。

リモートワークを行う従業員の(remote_worker_pc)と攻撃者が所持する端末(hacker_dev)はシステムの外にある。

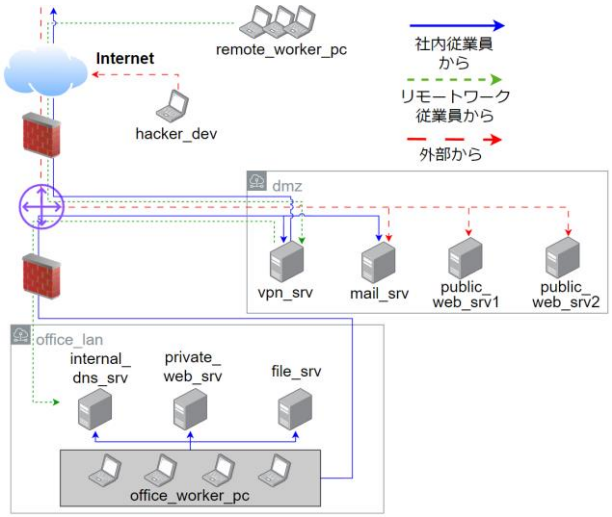

図 4　検証用シナリオに関するネットワーク構成図

追加の前提条件として、VPN サーバーと 3 つの Web サーバーが脆弱性を持っているとする。

Problem ファイルには、上記の内容を初期状態として記述する。そして、攻撃者がファイルサーバーの情報流出を達成した(achieved_TA0010 file_srv)を目標状態としてプランニングを実行した。

3.3.2 結果

Action のコストを全て 1 として行った 1 回目のプランニングの結果(プラン)を図 5 に示す。

```
   (t1566 hacker_dev office_worker_pc office_lan mail_srv
dmz mail1)
   (ta0002 office_worker_pc)
   (t1041 office_worker_pc file_srv)
; cost = 3 (unit cost)
```

図 5　1 回目のプランニングの結果

T1566 の Technique 名は「Phishing」であり、TA0002 の Tactics 名は「Execution」である。出力されたプランは、攻撃者が office_worker_pc に対してフィッシングメールを送信し制御を奪ってそこから file_srv の情報を流出するという内容である。

そして、1 回目に出力されたプランで用いられた全ての Action のコストを 1 増やした後にプランニングを行った。しかし、出力されたプランの内容は 1 回目と変わらず、Action の合計コストが 6 に増えたのみであった。

次に 1 回目のプランニングの後のコストの増分を 10 として行った 2 回目のプランニングの結果を図 6 に示す。T1210 は「Exploitation of Remote Services」、T1190 は「Exploit Public-Facing Application」という Technique 名で

ある。出力されたプランは、攻撃者が vpn_srv の脆弱性を悪用して office_lan 内に侵入して private_web_srv に対して脆弱性を悪用した攻撃で制御を奪ってそこから file_srv の情報を流出するという内容である。

```
   (t1210 hacker_dev vpn_srv dmz vpn1 office_lan)
   (t1190 hacker_dev private_web_srv office_lan prv_web1)
   (ta0002 private_web_srv)
   (t1041 private_web_srv file_srv)
; cost = 24 (general cost)
```

図 6　2 回目のプランニングの結果

2 回目に出力されたプランで用いられた全ての Action のコストを再度 10 増やして行ったプランニングで出力されたプランの内容は 1 回目と変わらず、Action の合計コストが 53 に増えたのみであった。

結果として、2 つのプラン内容が出力された。

3.4 評価

検証において出力された 2 つのプランに関わらないホストは internal_dns_srv、public_web_srv1、public_web_srv2 の 3 つである。これらのホストを利用した攻撃は実現可能性のない攻撃と判断できる。チェック項目の例として「5 分でできる！情報セキュリティ自社診断」の「パソコンやスマホなど情報機器の OS やソフトウェアは常に最新の状態にしていますか？」を考えた場合は、プランに関わらないホストをこのチェック項目の適用対象から外すことで、OS やソフトウェアが最新でない場合に起こりうる攻撃の過大評価を防ぐことができる。

このように、今回の検証からは file_srv の情報を流出する攻撃の対策を講じる上で全てのホストに対してチェックリストを適用した際に生じる過大評価を防ぐこと及び提案手法の有効性を確認できた。

また、この提案手法はプランニングによって攻撃者の攻撃パターンを全て導出できることが前提でチェックリストの適用対象を限定し攻撃の過大評価を防ぐ。しかし、コストの増分が 1 の時に全てのプランを出力できず増分を大きくしなければならなかった点から、コストの変化のみで全てのプランを出力させることは難しく安定性に欠けるといえる。

4. 考察

3.4 で述べたように、PDDL ファイルでコストを変化させるだけでは全ての攻撃パターンを導出することが難しい。プランニングは性質上、最適な 1 つのプランを求める。よって、既存のプランナーでは全ての解決策を求めるためには PDDL ファイルに変化を加えながらプランニングを繰り返す必要があった。しかし、多くのシステムを対象にしたとき、それぞれでプランニングの最中に PDDL ファイルを編集しなければならないという点は出力の再現性及び安定性に欠ける。そこで、最適な 1 つのプランにとどまらず全てのプランを導出するプランナーを作成することが望ましいと考える。

ゼロデイ攻撃に対して提案手法が対応しているかどうかという点については、PDDL の記述方式によって対応できると考える。MITRE ATT&CK は既存の攻撃手法に関する情報のみを記載しているが、PDDL ファイルにその内容を反映する際に「サービス X に関する脆弱性」といったように抽象的な記述の仕方をしている。よって、ゼロデイ脆弱性がサービス X に関する脆弱性である場合は、事前に提案手法を用いたリスクアセスメントを行えば「サービス X に関する脆弱性」がある場合として既に考慮された形で結果が出力されリスク対策の決定を補助する。

提案手法の汎用性については、MITRE ATT&CK などの要素を取り入れた PDDL ファイルの記述方式のフレームワークを作成する必要があると考える。3.3 では 1 つのケースでしか検証をしておらず使用した PDDL ファイルも対象ケースに特化した記述となっている部分もある。例えば、Web サーバーに関する Techniques は今回出力されたプランに含まれるもの以外にも多くある。しかし、今回のシステムと攻撃に関係がありそうな Techniques を判断した上で該当するもののみを PDDL ファイルに反映している。

この提案手法のあるべき形は、MITRE ATT&CK の情報を反映した Domain ファイルが 1 つ既に用意されており、リスクアセスメントの対象となるシステムの情報についてはそれぞれが Problem ファイルに記述し、プランニングを実行するというものである。2024 年 5 月現在、Techniques の数は Sub-techniques を含めて 400 以上にのぼる。これらを全て整合性が取れた形で Predicates と Action を定義した Domain ファイル、及び誰もが Problem ファイルを記述しやすい PDDL ファイルの記述方式を作成する必要がある。

5. おわりに

本論文では、IT システムのリスクアセスメントにおけるチェックリスト利用法の課題に対処するために、MITRE ATT&CK などの要素を取り入れたプランニングによるリスクアセスメント手法を提案した。この手法により、攻撃者が目的を達成するための現実的な攻撃シナリオのみを評価対象とすることが可能となり、従来のチェックリスト利用法における実現可能性の低い攻撃の過大評価を防ぐことを示した。

今後は、提案手法の前提となる全てのプランを出力可能であるプランナーを作成し、MITRE ATT&CK などの要素を取り入れたプランニングが様々な種類や複雑なシステムを対象とした場合も成功する PDDL ファイルの記述方式を考案し実装していく。

参考文献

[1] “MITRE ATT&CK®”, https://attack.mitre.org, (参照 2024-05-30)
[2] ” ISO - ISO 31000 — Risk management”, https://www.iso.org/iso-31000-risk-management.html, (参照 2024-05-30)
[3] 佐々木良一(編著), “IT リスク学：情報セキュリティを超えて”, 共立出版, 2013, 144-147p
[4] “5 分でできる！情報セキュリティ自社診断_情報セキュリティ_ IPA 独立行政法人 情報処理推進機構”, https://www.ipa.go.jp/security/guide/sme/5minutes.html, (参照 2024-05-30)
[5] “中小企業の情報セキュリティ対策ガイドライン_情報セキュリティ_IPA 独立行政法人 情報処理推進機構”, https://www.ipa.go.jp/security/guide/sme/about.html, (参照 2024-05-30)
[6] “HomePage - Fast Downward Homepage”, https://www.fast-downward.org, (参照 2024-05-30)
[7] LaValle, Steven M. “Planning algorithms”, Cambridge university press, 2006