proofpoint.

人とAIを中心とした脅威を防ぐための CISO向けガイド

主な機能

「人」やAIを中心とした脅威から組織を保護するために必要な5つの機能

1. 包括的な脅威の可視性とリスクに関する知見
2. メールやその他のチャネルをカバーする自動化された脅威対策
3. ビジネス上の信頼されたコミュニケーションのセキュリティ
4. 従業員向けのガイダンス
5. アカウント乗っ取り対策

概要

攻撃者は依然として、金銭的利益を得るために、データの持ち出しやビジネス上のコミュニケーションの悪用に力を注いでいます。また、こうした脅威の数は増加の一途をたどる一方で、攻撃手法自体はほとんど変わっていません。フィッシング、マルウェア、ランサムウェア、ビジネスメール詐欺（BEC）およびソーシャル エンジニアリングは、現在も人を標的とする攻撃の一般的な手法です。

新しい点は、AIによって、こうした従来型の攻撃手法がさらに高度化していることです。攻撃者は、大規模言語モデルを使用して、高度にパーソナライズされたフィッシングの誘い文句を作成し、攻撃チェーンの80～90%を自動化し、前例のない規模で多段階・多チャネルのキャンペーンを展開しています。プルーフポイントでは、2025年だけで顧客を狙ったメールの脅威が94%増加したことを確認しています。AIは、メールに埋め込まれた隠れた指示を通じて企業のAIアシスタントを武器化するプロンプト インジェクション攻撃など、まったく新しい攻撃経路を生み出しています。

このガイドでは、メールやその他のチャネルにおいて、人とAIを中心としたあらゆる脅威に対し強力な防御を築くために必要な重要な機能について掘り下げます。 また、組織に適した、セキュリティ プラットフォームを選ぶにあたり、検討すべき事項も提案します。

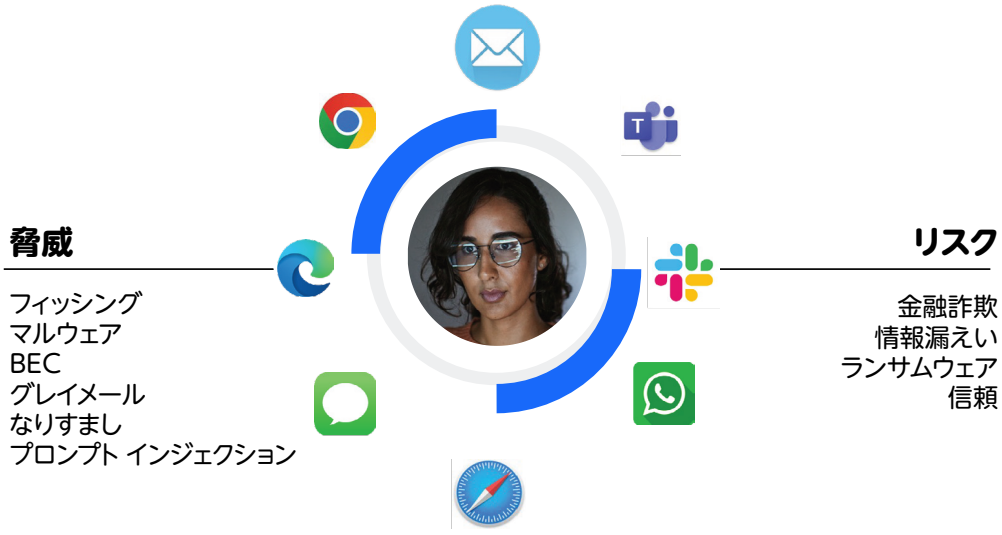

図1:デジタルワークスペースにおける脅威とリスク

1:包括的な脅威の可視性とリスクインサイト

人や AI を中心とした脅威を阻止するには、どのユーザーが、どのように標的にされているかを把握する必要があります。これにより、特にリスクの高いユーザーに適応型セキュリティ制御を適用できます。

メールおよびデジタルチャネル全体にわたって脅威を包括的に可視化すれば、脆弱性の全体像を把握できます。

可視化すべきポイント

- **誰が標的にされているか**：ユーザーがどのような脅威に直面しているか、攻撃者とやり取りの有無など
- **フォレンジックの詳細**：攻撃者、脅威ファミリー、感染したユーザー、攻撃手法とテーマ、攻撃キャンペーンの目標など
- **リスクにさらされているユーザー**：組織にリスクをもたらす人と理由を特定
- **信頼関係に基づくビジネスコミュニケーション内の脅威**：ブランドの評判を低下させるような類似ドメインや偽装されたドメイン、Web サイトなど
- **行動の変化と脅威インテリジェンス**：サプライヤーまたは信頼できるサードパーティが侵害されている可能性があることを示す兆候
- **不審なアクティビティ**：アカウント乗っ取りの可能性を示す兆候

AI を活用したプラットフォームは、これらの領域にわたるシグナルを関連付け、関係グラフを使用して正常なコミュニケーション パターンのベースラインを設定し、言語モデルを使用してメッセージの意図を解釈し、脅威インテリジェンスを使用して攻撃者の行動をコンテキスト化します。これにより、手動分析だけでは得られない、より深く、より実用的なリスクに関する知見を得ることができます。

可視性は初期導入時だけでなく、その後も継続して維持される必要があります。これにより、攻撃の変化に合わせて、保護レベルを継続的に調整できます。

2:メールとその先を保護する自動化された脅威対策

脅威状況は日々変化します。残念ながら、組織は多くの場合、こうした脅威に対応できるセキュリティの人材やリソースが十分ではありません。その結果、チームにとって各セキュリティイベントを調査する時間がないことも珍しくありません。さらに、こうしたイベントによる被害額も増加傾向にあります。

そのため、生産性に影響を及ぼすことなく、脅威を正確かつ効率的に検知し、阻止できるソリューションが求められます。AI により生成され、AI により配信される脅威の増加に伴い、組織の検知能力も AI 対応であることが求められています。

自動化すべきポイント

- **脅威を受信前に阻止**：ユーザーの受信トレイに届く前に99.999%以上の有効性で阻止
- **AI生成の脅威を検知してブロック**：AIにより作成されたBECメッセージ、AIによりカスタマイズされたフィッシング、そしてMicrosoft Copilot などのAIアシスタントを狙った隠れたプロンプトインジェクション攻撃など
- **内部送信メールの行動パターンを分析**：AIおよび機械学習モデルを活用した脅威インテリジェンス技術によってラテラル フィッシングの検知
- **悪意のあるURLをリアルタイムで調査してブロック**：メール、またはメッセージング/コラボレーション プラットフォーム経由でのユーザー到達を阻止
- **アイデンティティプロバイダー(IdP)のアカウントの侵害検知と対応**
- **配信前に不審なQRコードを分析**：AIを活用したコンピュータビジョンとセマンティック分析を使用したサンドボックス処理
- **不審なメッセージに警告タグの挿入**

攻撃者が初期アクセスに成功した場合、こうした脅威を迅速に検知して対応できることが重要です。こうすることで大規模な侵害になりそうなものをマイナー インシデントに抑えることができます。

488万ドル

フィッシングまたは
BEC攻撃による
データ侵害の平均被害額[1]

1.IBM、Cost of a Data Breach Report(データ侵害にかかる費用のレポート)、2024年

3:ビジネス上の信頼されたコミュニケーションのセキュリティ

デジタルコミュニケーションは組織の生命線です。そのため、攻撃者が信頼関係に基づくコミュニケーションを狙うのは当然です。受信者が信頼できる相手とやり取りしていると信じ込ませることができれば、BEC、フィッシング、ランサムウェアなどの攻撃は、成功率が高くなります。

攻撃者は、人をだます可能性を高めるために、多様ななりすましの手法を用います。AIによって、なりすましの効果は飛躍的に高まりました。攻撃者は、経営者の口調や文体を模倣した、文脈に即した完成度の高いメッセージを数秒で生成できるようになりました。これらを阻止するには複数の保護レイヤーが必要です。

ソリューション選定のポイント

- **メール認証**: ユーザー作成メールとアプリケーション生成メールにメール認証を適用
- **セキュアな専用環境**: アプリケーションが生成するトランザクション メールをリレーするための環境を提供
- **DMARC実装のサポート**: メール認証の効果を最大化し、完全なDMARCコンプライアンスを確保
- **類似ドメイン対策**: 類似ドメインを検知し、ブロックや閉鎖措置を支援
- **侵害されたサプライヤー アカウントの監視**: 振る舞いAIと脅威インテリジェンスのほか、侵害されたサプライヤー アカウント経由の脅威に対し、自動アクションで防御

ビジネス上の信頼されたコミュニケーションを保護できれば、従業員だけでなく、ビジネスパートナーや顧客も保護できます。

4:従業員向けのガイダンス

テクノロジーで脅威の99%を阻止できたとしても、残りの1%が大規模なインシデントを招く可能性もあります。ここで鍵を握るのが、人の行動です。攻撃者は通常、悪意のあるキャンペーンを成功させるために、従業員の何らかの関与を必要とします。

懸念すべきは攻撃だけではありません。ユーザーは多くの場合、利便性を優先して、組織のセキュリティを犠牲にしてしまうことがあります。Proofpoint 2024 State of the Phishレポートによると、

- 従業員の71%が、パスワードの使いまわしや不明なリンクのクリックなど、リスクのある行動を取ったことを認めています。
- これらの従業員の96%が、自身の行動にリスクが伴うことを承知の上で行っていました。

AIツールが日常のワークフローに組み込まれるにつれて、従業員は未承認のAIアプリケーションと機密データを共有したり、AIアシスタントとやり取りすることで隠れたプロンプト インジェクションを意図せず作動させるなど、新たなリスクに直面しています。

攻撃に不注意なユーザー行動が重なると、侵害の確率はさらに高くなります。そこでユーザーの教育が必要になります。

ソリューション選定のポイント

- **脅威データの活用**: 最も標的にされているユーザーや高リスクのユーザーを特定
- **リスクベース教育のユーザー提供**: 実際に自組織を標的とする脅威に即した実例を活用
- **行動変容にフォーカス**: 年1回のセキュリティトレーニングを「こなす」だけでは不十分
- **従業員の意欲向上**: 個人のリスクスコアと、個人が組織のセキュリティ ポスチャ(態勢)に与える影響度を可視化
- **有効性の評価**: 戦略を洗練できる重要なレポートを提供
- **トレーニングコンテンツでAIに関連したリスクに対処**: 生成AIツールの安全な利用方法やAI生成のソーシャル エンジニアリング攻撃の認識方法をカバー

強力なテクノロジーと人間の注意力を組み合わせることが、「人」を中心とした脅威を阻止する上で重要です。誰もが、業務運営を安全に保つうえで重要な役割を担っています。

71%

従業員の71%が、パスワードの使いまわしや不明なリンクのクリックなど、リスクのある行動を取ったことを認めています。[2]

2.プルーフポイント、2024 State of the Phishレポート、2024年

5:アカウント乗っ取り対策

プルーフポイントのデータによると、組織の99%がアカウント乗っ取り(ATO)の試行に定期的に直面しています。これらの攻撃は、ID窃取の一形態として、サイバー犯罪者がオンライン アカウントにアクセスする、または「乗っ取る」ものです。Microsoft Entra ID、Google、Oktaなどのクラウドベースのアイデンティティ プロバイダー(Idp)が最も標的にされていることは何ら驚きではありません。これらのアカウントは、ユーザーが利用するエンタープライズ アプリケーションへのシングルサインオン(SSO)ゲートウェイとして機能します。

そして注意が必要なのはアカウントだけではありません。サイバー犯罪者は信頼されたビジネスパートナーのアカウントを侵害して偵察を行い、さらなる攻撃を行います。 これらの侵害されたアカウントは、多段階攻撃の侵入点となり、組織のエコシステム全体に拡散し、機密データを盗み、不正な取引を行い、混乱を引き起こします。

AIと機械学習は、大規模なビジネスコミュニケーションを監視し、対応を自動化するために不可欠です。振る舞いAIモデルは、異常なログインパターン、異常なメール送信行動、またはコミュニケーション関係の変化など、ルールベースのシステムでは見逃すような、アカウント侵害の微妙な兆候を検出できます。

ソリューション選定のポイント

- **すべてのアカウントを継続的に監視**: Microsoft Entra ID、Google、Oktaなどのクラウドベースのidpのアカウントを監視
- **脅威インテリジェンスを使用**: 振る舞いデータや機械学習と組み合わせて侵害されたアカウントを検知
- **アカウント乗っ取り攻撃から防御**: 乗っ取られたアカウントの65%が多要素認証(MFA)を使用[4]
- **調査を迅速化**: ATO後のアクティビティに関する一元化されたビューを提供
- **対応機能を自動化**: アカウントの停止、パスワードリセットの強制実行、メールボックスのルールやMFA設定への悪意のある変更の修復
- **不審なサードパーティ アプリケーション**を削除(ATO後のクリーンアップ)

ATOは金銭的被害だけでなく、ブランドの風評被害ももたらす可能性があります。リスクを減らすには強力な保護が欠かせません。

ツールの分断を避ける

メールやその他のチャネルの防御を構築するにあたり、複数の専門ベンダーのポイントソリューションを組み合わせることが、当然の解決法に思えるかもしれません。確かに、専門ベンダーは、特定のタイプの攻撃への対処に長けているように思えます。しかし、このサイロ化されたアプローチではいくつもの欠点が見られます。

まず、セキュリティ上の盲点が生まれます。ツールがシームレスに統合されていなければ、セキュリティチームにとって、セキュリティ環境全体を可視化するのが難しくなります。その結果、脅威が検知されないリスクが高まるばかりか、インシデント レスポンスにも遅れが生じます。

断片化されたツールは、今日の脅威を阻止するために必要なアンサンブルAIアプローチを提供することもできません。言語モデル、コンピュータビジョン、振る舞い分析、および脅威インテリジェンスが連携し、コンテキストを共有することで、洗練されたAI生成の攻撃を検出する必要があります。

複数のセキュリティツールを管理することは、チームにとって時間がかかり、非効率です。加えて、サイロ化された制御ポイント間でデータを相関させる必要があります。また、これらのプラットフォームが発生させる膨大な数のアラートは、アラート疲労を引き起こし、脅威を見逃す原因となります 。これらすべてを考慮すると、運用コストも高くなります。

そこで、より効果的なアプローチをおすすめします。「人」を中心としたあらゆる脅威に対処する、事前統合済みの包括的なセキュリティ プラットフォームを採用しましょう。信頼できるパートナー1社を利用すれば、管理を効率化できるだけでなく、財務面でのメリットも得られます。

99%

アカウント乗っ取り(ATO)の試行に定期的に直面する組織[3]

3. プルーフポイント調べ

4. 同上

結論

「人」を中心とした包括的なセキュリティ戦略を採用すれば、組織をさまざまな脅威から保護することができます。こうした目標に向けての最初の一歩は、適切なソリューション選びです。メール、コラボレーションツール、メッセージングプラットフォーム、クラウドアプリケーションにわたって脅威インテリジェンスを統合できるソリューションを選びましょう。リスクのあるユーザー行動について重要な知見を提供し、セキュリティ文化を向上できるものが求められます。

現在使用しているソリューションは、メールだけに対応したものでしょうか?また、分断化されたポイントソリューションを利用していませんか?これらが当てはまる場合、改善の余地があります。メールやその他のチャネルにおいて、「人」を中心としたあらゆる脅威に対し、組織のセキュリティが保護できているかを評価しましょう。

プルーフポイントで統合

Proofpoint Threat Protection Primeは事前統合済みの脅威プロテクション プラットフォームを提供し、完全なセキュリティを実現します。本ソリューションは、メールやデジタルチャネルなど、最新のワークスペースにおいて脅威をブロックします。さまざまな脅威を比類のない精度で検知しながら、こうした脅威から保護します。また、言語モデル、機械学習、コンピュータビジョン、関係グラフ、脅威インテリジェンスを含むAIエンジンのアンサンブルであるProofpoint Nexus AIプラットフォームを活用し、従来の脅威とAI生成の脅威の両方に対して99.999%の有効性を提供します。

また、人的リスクに関する詳細な知見を提供し、ユーザーのレジリエンスを強化します。侵害されたユーザーアカウントやサプライヤーアカウントを悪用した脅威からも防御し、ビジネス上の信頼されたコミュニケーションの安全を守ります。

プルーフポイントは、組織にとって最大の資産であり、最大のリスクである従業員を保護するための適応型アプローチと共に、唯一の最新セキュリティ アーキテクチャを提供しています。こうした理由から、プルーフポイントは、Fortune 100企業の80社以上含む、あらゆる規模の270万以上のお客様にご利用いただいています。

proofpoint.

プルーフポイント プラットフォームの詳細はこちら →