Vol.2018-CSEC-82 No.20
Vol.2018-SPT-29 No.20
2018/7/25

セーフティ機能のセキュリティ脅威に対する効果の分析

大久保 隆夫[1,a)]

概要：セーフティとセキュリティの問題で重要視されるのは，セキュリティ脅威 (攻撃など) によるセーフティへの影響の問題である．しかし，現実にはセーフティとセキュリティは互いに影響しあっているため，セーフティに対する脅威がセキュリティに影響を及ぼす (例えば，情報漏洩や改ざんなど) 可能性も考慮しなければならない．本研究では，セーフティの脅威への既存対策が，セキュリティ脅威に対してどの程度有効に機能するのか，分析を行った．また，セキュリティ，セーフティ双方が要件となりうる自動車の自動運転の例にこの結果を適用し，妥当性を確認した．

Impact Analysis of Safety Functions against Security Threats

Okubo Takao[1,a)]

1. はじめに

セーフティとセキュリティの問題で重要視されるのは，セキュリティ脅威 (攻撃など) によるセーフティへの影響の問題である．DEFCON および BlackHat で発表された車載システムへの攻撃の事例 [1][2][3] はいずれも，セキュリティ脅威がセーフティに及ぼす影響について指摘したものである．しかし，現実にはセーフティとセキュリティは互いに影響しあっているため，セーフティに対する脅威がセキュリティに影響を及ぼす (例えば，情報漏洩や改ざんなど) 可能性も考慮しなければならない．

本研究では，セーフティの脅威への既存対策が，セキュリティ脅威に対してどの程度有効に機能するのかを，分析により確認した．既存のセーフティ対策は，文献 [7] をもとに列挙した．分析の結果，効果は，不明なものや該当する脅威のないものを除き，「ある程度あり」「ほぼなし」「なし」「逆効果」に分かれた．

2. 研究の背景，関連研究

IoT 機器を踏み台に悪用するマルウェア「Mirai」の流行や，BlackHat/DEFCON において自動車や IoT 機器がしばしばハッキング対象としてとりあげられるなど，IoT や組み込み機器に対するセキュリティ脅威が高まっている．そのため，開発者は従来のセーフティ系ハザード，リスク分析の他に，セキュリティのリスク分析も行うことが求められている．セーフティ系リスクとセキュリティ系リスクには，共通するものもあるため，セーフティ系分析とセキュリティ分析をいずれか一方だけ行うことにより，もう一方の分析も網羅できるのであれば，システム，機器開発において効率化，省力化が期待できる．

そのため，筆者らはセキュリティとセーフティを統合的に分析可能かの検討を進めている．筆者らは，リスク評価についてセーフティ，セキュリティ双方の手法を調査し，リスクにおける事象発生確立の扱いや評価の主観性において，両者の評価手法には相違があることを明らかにした [4]．現在は，対策面に焦点をあて，セキュリティ対策，セーフティ対策の，お互いの脅威に対する有効性の検証を行っている．セキュリティ対策のセーフティ脅威への有効性は，機器に対するハッキング事例の分析の中で検討されていることが多い．一方，セーフティ対策のセキュリティ脅威に対する有効性は，あまり検討されていない．そこで，筆者らはセーフティ対策のセキュリティ脅威に対する有効性を分析することにした．

セキュリティとセーフティを統合する分析手法としては，STAMP/STPA の拡張である STPA-sec[5] および STPA-

[1] 情報セキュリティ大学院大学
IISEC, Tsuruyamachi 2-14-1, Kanagawa-ku, Yokohama 221–0835, Japan
[a)] okubo@iisec.ac.jp

safesec[6] がある．STPA-sec は，ハザード分析である STPA において安全でない (unsafe) コントロールアクションの識別に加え，セキュアでない (unsecure) コントロールアクションの識別を行うことでセーフティ分析とセキュリティ分析の統合をはかっているが，セキュリティ分析を補足的に追加しているだけで，セキュリティとセーフティの相互作用ういては言及されていない．また，STPA-safesec では，セキュリティ脅威がセーフティに与える影響を考慮したコントロールアクションの識別を行っている．しかし，本稿で対象にしているセーフティ機能のセキュリティ脅威への影響については，言及がない．

3. 分析方法

本稿では，セーフティの脅威への既存対策が，セキュリティ脅威に対してどの程度有効に機能するのかを，分析により確認した．手順を以下に示す．

(1) 対象となるセーフティ対策の列挙
セーフティ対策は，文献 [7] にある機能安全の対策分類をもとに列挙した．

(2) セキュリティ脅威の列挙
これらの対策に対し，脅威モデリングで用いるガイドワード STRIDE[8] の 6 種類の脅威 (s:なりすまし，T:改竄，R:否認，I:情報漏洩，D:サービス妨害，E: 特権昇格) を採用した．また，セーフティを考慮する場合，BlackHat などにおける機器の不正操作，乗っ取りを表すのに STRIDE では不十分と考え，新たに「不正操作(起動，停止含む)」「悪用」，「のっとり」を追加した．これらの脅威ガイドワードを対策ごとに列挙した．

(3) 効果の評価
次に，各対策がセキュリティ脅威に対しどれだけ効果があるかを評価する．効果の評価は筆者が主観評価により「あり」「ある程度あり」「ほぼなし」「なし」「逆効果」の 5 段階で評価した．また，セーフティ対策の導入自体が新たなセキュリティ脅威になっている場合はそれも記した．また，セキュリティ脅威との関連性が不明だったものについては「不明」と記した．

4. 分析結果

分析の結果を，表 1〜3 に示す．

分析の結果，効果は，不明なものや該当する脅威のないものを除き，「ある程度あり」「ほぼなし」「なし」「逆効果」に分かれた．それぞれ，代表的な例を以下に示す．

- ある程度あり
  例:SW(ソフトウェア) 多様性 (ID21)
  ある機器または類似の部品が一斉に同じ原因で故障すると，ソフトウェアに冗長性を持たせていても，冗長性の意味がなくなってしまうため，ソフトウェアに多様性を持たせることで，全体の故障する可能性を小さくできる機能．この機能は，サービス妨害やソフトウェアに対する不正操作や乗っ取りに対しては，ある種の攻撃が成功したとしても，別の種類のソフトウェアに対しては同種の攻撃が成功しないこともあるため，多様性はある程度効果があると言える．
- ほぼなし
  例:SW 冗長性 (ID20)
  あるソフトウェア部品が故障した場合に備え，バックアップ的に動作するソフトウェアを用意するなどの機能．冗長性を持たせる場合に，単純に同種のソフトウェアを用いている場合，そのソフトウェアに対し有効な攻撃によって，冗長なソフトウェアも被害を受け，冗長性が機能しなくなってしまうおそれがある．
- なし
  例:一つの機械を複数の制御器で起動できる場合は，稼動時はそのうちの一つが有効となるような制御設計 (ID7)
  なりすましによる不正な稼動の脅威に対し，攻撃者にとって，複数の制御器のどれかが有効であれば，なりすましによる稼動が有効となるため，効果はない．
- 逆効果
  例:SW 故障に対する検知 (ID15)
  SW が故障した際，即座に検知できる機能．これにより，故障に対する迅速な対処を行うことができる．この機能に対し，故障情報の改竄を想定する．まず，故障があるのにないように改竄された場合，その検出は困難になるため攻撃に対する効果はない．一方，故障がないのにあるように改竄させる場合については，この改竄を悪用して何度も偽の故障情報を送りつけ，サービス妨害を起こさせる危険がある．この場合，セキュリティ脅威に対する効果としては逆効果になる．

5. 自動運転に対する適用結果

本分析結果の妥当性について確認するため，現実に機能安全対策実施されている，または実施が計画されているシステム事例に適用を行った．対象のシステムとしては，近年注目されており，前述のハッキングの事例からセキュリティ，セーフティ双方の要件が必要となる自動車の自動運転システムを採用した．自動運転に対するセーフティ対策は，文献 [9] に記載されている．挙げられている機能安全対策を表 4 に示す．

- 自動ブレーキ
  自動ブレーキは，自動運転技術の先駆けとして実現された技術である．自動ブレーキは，障害物との距離などの周辺状況をセンサーから受けとり，危険な状況を判断して自動的にブレーキを作動させる．この自動ブレーキ機能に対しては，センサーからの信号を改竄する脅威が想定される．これについては，表のと同様の

表 1 分析結果 (1)

対策 ID	機能安全対策分類	機能安全対策	セキュリティ脅威への有効性	セキュリティ脅威	新たなセキュリティ脅威
1	制御システムへの本質的安全設計の適用	機構運動の起動または停止	なりすましによる起動または停止	なし	
			起動信号，停止信号の改竄	ある程度あり	
			起動または停止事実の否認	該当脅威なし	
			起動，停止情報の漏洩	なし	
			起動または停止に対する妨害	ある程度あり	
			特権昇格による起動，停止	なし	
			不正な起動，停止	ある程度あり	サービス妨害に悪用されるリスク
2		動力中断後の再起動防止	なりすましによる再起動	ある程度あり	
			再起動信号の改竄	ある程度あり	
			再起動の否認	該当脅威なし	
			再起動情報の漏洩	なし	
			再起動に対する妨害	ある程度あり	
			特権昇格による再起動	なし	
			不正な再起動	ある程度あり	
3		動力供給の中断	なりすまし	ある程度あり	
			動力供給信号の改竄	ある程度あり	
			動力供給の否認	該当脅威なし	
			動力供給情報の漏洩	なし	
			運転に対する妨害	ある程度あり	
			特権昇格による動力供給の中断	なし	
			不正な動力供給の中断	なし	
4		自己監視の使用	なりすまし	ある程度あり	
			情報の改竄	ある程度あり	
			否認	ある程度あり	
			情報の漏洩	ある程度あり	
			運転等に対する妨害	ある程度あり	
			特権昇格	ある程度あり	
			不正操作	ある程度あり	
5		人間工学の原則に従った設計	不明		
6		停止制御装置は起動制御装置の近くに配置	なりすまし	ある程度あり	
			情報の改竄	なし	
			否認	なし	
			情報の漏洩	なし	
			運転等に対する妨害	ある程度あり	攻撃者が停止
			特権昇格	ある程度あり	制御装置を悪用
			不正操作	ある程度あり	
7		稼動時は複数の制御器の一つが有効となるような制御設計	なりすまし	ある程度あり	
			情報の改竄	なし	
			否認	なし	
			情報の漏洩	なし	
			運転等に対する妨害	ある程度あり	攻撃者が停止
			特権昇格	ある程度あり	制御装置を悪用
			不正操作	ある程度あり	
8		無線通信による制御の場合は，制御信号が受信されない場合は自動停止となる設計	なりすましによる遠隔制御	ある程度あり	
			無線通信制御命令の改竄	ある程度あり	
			否認	なし	
			制御情報の漏洩	なし	
			制御に対する妨害	逆効果	
			特権昇格	ある程度あり	
			不正操作	ある程度あり	

表 2 分析結果 (2)

対策 ID	機能安全対策分類	機能安全対策	セキュリティ脅威への有効性	セキュリティ脅威	新たなセキュリティ脅威
9		制御モード及び運転モードの選択	なりすましによる運転	なし	
			改竄	該当脅威なし	
			否認	該当脅威なし	
			情報の漏洩	該当脅威なし	
			いずれかのモードに対する妨害	ある程度あり	
			特権昇格	ある程度あり	
			自動運転に対する乗っ取り	ある程度あり	
10	幾何学的及び物理的要素に関する配慮		不明		
11	機械設計に関する一般的技術知識の考慮		不明		
12	機械的結合の安全原則		不明		
13	人間工学原則の遵守		不明		
14	安全機能の故障の確率の最小化		不明		
15	空圧及び液圧設備の危険源の防止	ソフトウェア (SW) 自己診断機能	なりすまし	該当脅威なし	
			故障情報の改竄	なし/逆効果	
			否認	該当脅威なし	
			情報の漏洩	該当脅威なし	
			サービス妨害	該当脅威なし	
			特権昇格	該当脅威なし	
			SW の不正操作，乗っ取り	ある程度あり	
16	電気的危険源の防止	HW，センサ，アクチュエータの監視	なりすまし	なし	監視機能の無効化や改ざん
			センサー情報等の改竄	ある程度あり	
			否認	該当脅威なし	
			センサー情報等の漏洩	該当脅威なし	
			サービス妨害	ある程度あり	
			特権昇格	該当脅威なし	
			センサー，機器ののっとり	ある程度あり	
			リプレイ攻撃など		
17		安全機能試験	不明		
18		SW の容量と応答時間性能	不明		
19		SW 安全妥当性確認計画	不明		
20		SW 冗長性	なりすまし	該当脅威なし	
			改竄	該当脅威なし	
			否認	該当脅威なし	
			情報漏洩	該当脅威なし	
			サービス妨害	ほぼなし	
			特権昇格	ほぼなし	
			SW の不正操作，乗っ取り	ほぼなし	
21		SW 多様性	なりすまし	該当脅威なし	
			改竄	該当脅威なし	
			否認	該当脅威なし	
			情報漏洩	該当脅威なし	
			サービス妨害	ある程度あり	
			特権昇格	ある程度あり	
			SW の不正操作，乗っ取り	ある程度あり	
22		ガード，センサなどの保護装置	不明		
23		インターロック装置	不明		
24		両手操作制御装置	不明		

表 3 分析結果 (3)

対策 ID	機能安全対策分類	機能安全対策	セキュリティ脅威への有効性	セキュリティ脅威	新たなセキュリティ脅威
25		イネーブル装置	遠隔からののっとりなど	ある程度あり	ボタン押下を偽造する攻撃による起動など
26		ライトカーテン	不明		
27		圧力マット	不明		
28		非常停止	不明		
29		動力源の遮断装置	不明		
30	または制限装置	蓄熱エネルギーの消散	不明		
31	要素を手で動かすための手段	非常停止後に特定の	不明		

表 4 自動運転の機能安全対策とセキュリティ脅威に対する効果

機能安全対策分類	機能安全対策	説明	セキュリティ脅威	セキュリティ脅威への有効性
自動ブレーキ			センサー情報の改竄 (危険でない内容に) センサー情報 (危険信号) の偽造	ほぼなし 逆効果
ドライバー異常時対応システム	押しボタン式検知 ハザードランプ，ブレーキランプ点灯 減速停止		ボタン押下時の改竄 信号の偽造 ランプの不正操作 (点灯) ランプの不正操作 (点灯，点滅させない) 不正操作 (減速させない)	ほぼなし 逆効果 ほぼなし ほぼなし ほぼなし
安全な受け渡し		システムが機能限界に陥る場合に 4 秒前にドライバーに警告	機能限界状態の発生 機能限界情報の改竄 設定情報 (時刻等) の改竄	ほぼなし ほぼなし ほぼなし
危険最小化防御		ドライバーが警告に応じない場合には車を安全に停止させること	安全停止信号の改竄 (無効化) 安全停止信号の偽造によるサービス妨害	なし 逆効果

効果となる．すなわち，センサー情報の改竄が，危険なのに危険でない方向に改竄された場合は，改竄により自動ブレーキは危険を検知できないため，効果は「ほぼなし」となる．一方，改竄が偽造，あるいは危険でない信号を危険に改竄された場合，サービス妨害あるいは急ブレーキによる事故の危険があり，自動ブレーキの本来の趣旨からみて，逆効果となる．

- 押しボタン式検知
  自動運転時に，異常を検知した運転者または乗客が押しボタンにより停車させる方式である．人が押しボタンを押してから，それが停止制御装置に伝わるまでに，信号の伝達が機械的に行われ，情報セキュリティ脅威の影響を受けない場合は，情報セキュリティ脅威はこの機能については該当しない，ただしその信号がサイバー攻撃により妨害することが可能な場合，上記自動ブレーキと同様のセキュリティ脅威が想定される．一つはボタンが押下されたのに，押下されていないように信号が改竄される場合で，この場合は表 2 の ID15「自己診断機能」と同様，ほぼ効果なしとなり，ボタンが押下されていないのに，押下されたように信号が偽造ないし改竄された場合は，逆効果となる．
- ハザードランプ，ブレーキランプ点灯
  異常を検知した際，ハザードランプ，ブレーキランプを点灯させ，周囲に知らせる方式．この方式については，ハザートランプ，ブレーキランプを不正に操作し，危険な状況時の点灯点滅を阻止したり，逆に安全時に点灯，点滅させる脅威に対しては，攻撃を受けた場合の効果はほぼないと認められる．
- 減速停止
  異常を検知した場合，自動的に減速，停止する方式．この方式についても，上記のハザードランプ，ブレーキランプ同様，減速停止機能を不正に操作する可能性

がある．

- 危険最小化防御
  「押しボタン式検知」の場合と同様，最小化を無効化ずるような改竄に対しては効果がなく，逆に必要のない最小化への改竄については，逆効果となる．

6. 議論

本稿の分析の結果は，セキュリティ対策に対する効果が「なし」であったり「逆効果」になった対策について，実際に効果がないかどうかは実装に依存するため，分析結果の確実性を保証するものではない．例えば，自動運転の「押しボタン式検知で，ボタン押下信号が純粋に電気信号として伝達され，サイバー攻撃の影響を受けない場合は，脅威そのものが存在しなくなる．分析においては，実装が不明であるため，最も脆弱な場合を想定し，効果は最も効果が期待できない場合の値を採用した．

また，セキュリティ脅威を想定しているにも関わらず，分析対象となりうる脅威のほとんどが直接可用性に関係するものか，完全性が対象でも間接的に可用性に影響するものとなってしまった．これは，機密性が個別のシステムに対するステークホルダの要求に依存するため，本稿において分析対象とした一般的なシステムにおいては，システム内で扱うデータのうちどのデータが機密性が必要になるかが想定しにくいことが原因と考えられる．

7. おわりに

本稿では，セキュリティ分析とセーフティ分析の統合の可能性を検討している中で，セキュリティ対策とセーフティ対策の共通化の見地から，セーフティ (機能安全) 向けの対策のセキュリティ脅威に対する対策としての有効性について分析を行った．その結果，既存のセーフティ対策機能に関しては，セキュリティ脅威に対してもある程度効果が期待できるものもあるが，ほとんど効果が期待できないものや，セキュリティ的には逆効果になるものもあることが分かった．したがって，組込みや制御システムにおいて，セキュリティ脅威を考慮しなければならない場合，既存のセーフティ対策では十分とは言えないため，別途セキュリティ分析，対策が必要になる．

今後は，対策検討，評価も含めてセーフティとセキュリティを統合的に扱う方式について研究を進める予定である．

謝辞 本研究は，株式会社富士通研究所との共同研究によるものです．ここに謝意を表します．

参考文献

[1] DEF CON 21, https://www.defcon.org/html/defcon-21/dc-21-index.html
[2] BlackHat USA 2014, https://www.blackhat.com/us-15/
[3] BlackHat USA 2017, https://www.blackhat.com/us-17/
[4] 大久保隆夫, セキュリティ，”セーフティのリスク評価手法に関する調査”,2018 年暗号と情報セキュリティシンポジウム (SCIS),2018.
[5] Ivo FriedbergMcLaughlin,Paul Smith,David Laverty,Sakir SezerKieran,STPA-SafeSec: Safety and security analysis for cyberphysical systems. Journal of Information Security and Applications, 2017.
[6] William Young, Nancy Leveson, “Systems thinking for safety and security”, Proceedings of the 29th Annual Computer Security Applications Conference, pp.1-8, 2013.
[7] 社団法人組込みシステム技術協会, 「組込み系技術者のための安全設計入門」, 電波新聞社, 2010.
[8] Adam Shostack, “Threat Modeling: Designing for Security”, Wiley, 2014.
[9] 高度情報通信ネットワーク社会推進戦略本部，「官民 ITS 構想・ロードマップ 2016」，2017.