2A2-1

制御通信における外れ値検出による侵入検知精度の比較

Comparison of Intrusion Detection Accuracy using Outlier Detection in Control System Communication

木内 舞[*1]
Mai Kiuchi

小野田 崇[*2]
Takashi Onoda

[*1] 財団法人 電力中央研究所
Central Research Institute of Electric Power Industry #1

[*2] 東京工業大学大学院 総合理工学研究科
Graduate School of Interdisciplinary Science and Engineering, Tokyo Institute of Technology #2

In this paper, we introduce outlier detection using SVM (Support Vector Machine) for intrusion detection in control system communication networks. SVMs are useful for classifying normal communication and intrusion attacks. In control systems, a large amount of normal communication data is available, but as there have been almost no cyber attacks, there is very little actual attack data. One class SVM and SVDD (Support Vector Domain Description) are two methods used for one class classification where only information of one of the classes is available. We applied these two methods to intrusion detection in an experimental control system network, and compared the differences in the classification.

1. はじめに

従来の監視制御システムにおける通信は，独自の通信方式を使用し，他のネットワークとの通信が限られていた．さらに，システムへの物理的なアクセスを難しくすることで，セキュリティが確保されてきた．しかし近年は監視制御システムでも，コスト削減等の理由により汎用の通信方式が採用されたり，運用性向上のために他のネットワークとの接続も検討されたりするようになってきた．このため，情報通信システムで問題とされているサイバーセキュリティと同様の問題が，監視制御システムでも問題になっている．

情報通信システムでは様々なセキュリティ対策が存在し，監視制御システムでも同様の対策の適用が検討されている [Kiuchi]．セキュリティ対策方法の一つである侵入検知システムは，通信ネットワークのパケットやシステムの動作を監視および解析し，不正な侵入を検出して通知する．これまでに，監視制御システムにおける侵入検知方法についての検討を行ってきた [Kiuchi]．

監視制御システムにおける侵入検知システムでは，正常な動作状態における正規通信のデータはあるが，サイバー攻撃がほとんど観測されていないため，不正通信のデータがない．このため，正規の通信パターンを規定し，規定したものからのずれによって侵入を検知する異常検知方式を採用することが考えられる．

異常検知方式の侵入検知システムを構築する方法の一つとして，サポートベクターマシン (SVM; Support Vector Machine) の適用が検討され，有効とされている [Chen]．SVM を侵入検知に適用した検討では，情報通信システムの通信パターンが評価に使用され，不正通信のデータについても知見があることが前提となっている．監視制御システムにおける侵入検知システムに SVM を適用するにあたっては，不正通信のデータがほとんど観測されていないことが問題となる．このため，一つのクラスのデータ例のみを使用して分類を行う手法の利用が妥当と考えられる．一クラスの分類問題に SVM を適用した手法としては，one class SVM [Schölkopf] や SVDD (Support Vector Domain Description) [Tax] が挙げられる．

情報通信システムでは，侵入検知システムへの one class SVM の適用についての検討も行われている [Zhang]．ただし，実ネットワークの通信において起こる回数が少ない通信を検出しており，その具体的な通信内容については検討されていない．また，文書検索の分野では，one class SVM と SVDD の適用が比較検討されている [Onoda]．

ここでは，監視制御システムの通信ネットワークにおける侵入検知に one class SVM および SVDD を適用することを検討した．それぞれの侵入検知結果の比較について報告する．また，模擬的な監視制御システムの通信データを使用し，正規通信と不正通信の判別を行えるかについて評価を行った．

2. 監視制御システムにおける侵入検知

セキュリティ対策は，最初のシステム構築時に行えばよいものではなく，継続的に行う必要がある．例えば，プログラムは人間が記述していることから，ミスや見逃しは避けられず，ソフトウェアの脆弱性を完全になくすことはほぼ不可能である．このため，セキュリティを保つには，脆弱性を発見した際に随時ソフトウェアをアップデートしていく必要がある．ソフトウェアのアップデートを行う際には，機器の再起動が必要となることも多い．

監視制御システムでは，監視制御対象として物理的に動作する機器に接続されていることや，24 時間稼働する必要があり，停止を許容しないケースも多い．このため，セキュリティ対策を施す際にシステムを停止させることが困難となり，セキュリティ対策を最新の状態に保つことが難しい場合が出てくる．そのような場合には，盗聴や改ざんを防ぐ暗号化，通信の正当性を確認する認証，システムへの侵入を防ぐファイアウォールの設置等といった単一の対策ではなく，複数の対策を組み合わせた多層的なセキュリティ対策を行うことでセキュリティを保つ．セキュリ

連絡先：木内舞，(財)電力中央研究所 システム技術研究所，
〒201-8511 東京都狛江市岩戸北 2-11-1
mai@criepi.denken.or.jp

ティ対策の一つとして，ネットワークの通信を監視し，不正な通信を検出して管理者に通知する侵入検知システムの導入が考えられる．

現状の侵入検知システムは，シグネチャ検知方式と異常検知方式の二種類に分けられる．シグネチャ検知方式では，侵入検知システムに既知の攻撃パターンや脆弱性を登録しておき，それに照らし合わせて不正な侵入を検知する．この方式は，既知の攻撃は正確に検知できるが，未知の侵入を検知できないという問題がある．もう一方の，異常検知方式では，正規の通信パターンを規定し，その規定からのずれによって侵入を検知する．この方式は未知の侵入を検知できるが，正規の通信パターンの規定が難しい場合も多い．監視制御システムの侵入検知システムでは，正常通信のデータは大量にあるが，不正通信のデータがほとんど観測されていない．このため，異常検知方式を採用することが考えられる．

異常検知方式の侵入検知システムを構築する方法の一つとして，SVM の適用が検討され，有効とされている．監視制御システムへの適用の際には，不正通信のデータがほとんど観測されていないことが問題となる．このことから，一つのクラスのデータ例のみを使用して分類を行う手法を利用することが妥当と考えられる．本検討では，one class SVM および SVDD の二つの方式の適用を検討した．

3. One class SVM と SVDD

3.1 One class SVM

今回検討した一つの手法である one class SVM は，SVM を一クラスの分類問題に適用したものである [Schölkopf]．入力データを非線形写像により高次元特徴空間上に写像し，学習データの多くが判別超平面を挟んで特徴空間上の原点との反対側にあることを維持しながら，超平面と原点のマージンが最大になるように超平面分離する．

入力するデータを $\mathbf{x}$ とすると，識別関数 f について，$f(\mathbf{x}) = +1$ の領域に学習データのほとんどが存在し，その他では $f(\mathbf{x}) = -1$ となるようにする．

$$f(\mathbf{x}) = \mathrm{sgn}\left(\left(\mathbf{w}\cdot\boldsymbol{\Phi}(\mathbf{x})\right) - \rho\right) \tag{1}$$

ここで，入力データ $\mathbf{x}_1, \ldots, \mathbf{x}_l \in \boldsymbol{X}$ とし，l個のデータが一つのクラス X に所属するものとする．$\boldsymbol{\Phi}: X \to H$ は高次元空間への写像で，簡単なカーネルで内積を計算できるものとする．

$$k(\mathbf{x},\mathbf{y}) = \left(\boldsymbol{\Phi}(\mathbf{x})\cdot\boldsymbol{\Phi}(\mathbf{y})\right) \tag{2}$$

なお，ここでは線形カーネルを使用した．

$$k(\mathbf{x},\mathbf{y}) = \mathbf{x}\cdot\mathbf{y} \tag{3}$$

識別関数の条件を満たすためには，次の二次計画問題を解くことになる．

$$\min_{\mathbf{w},\xi,\rho} \frac{1}{2}\|\mathbf{w}\|^2 + \frac{1}{\nu l}\sum_i \xi_i - \rho \tag{4}$$

制約条件：

$$\left(\mathbf{w}\cdot\boldsymbol{\Phi}(\mathbf{x}_i)\right) \geq \rho - \xi_i\,, \qquad \xi_i \geq 0\,, \qquad 0 < \nu \leq 1 \tag{5}$$

パラメータ ν はマージン誤差率で，学習データにおける外れ値の割合の上限値となる．また，学習データにおけるサポートベクトルの割合の下限値となる．

ここで，$\alpha_i, \beta_i \geq 0$ として，次のラグランジュ関数 L を導入する．

$$\begin{aligned} L(\mathbf{w},\boldsymbol{\xi},\rho,\boldsymbol{\alpha},\boldsymbol{\beta}) = {} & \frac{1}{2}\|\mathbf{w}\|^2 + \frac{1}{\nu l}\sum_i \xi_i - \rho \\ & - \sum_i \alpha_i\left((\mathbf{w}\cdot\mathbf{x}_i) - \rho + \xi_i\right) - \sum_i \beta_i\xi_i \end{aligned} \tag{6}$$

変数 $\mathbf{w}, \xi_i, \rho$ について導関数を 0 とすると，次が成立する．

$$\mathbf{w} = \sum_i \alpha_i \boldsymbol{\Phi}(\mathbf{x}_i) \tag{7}$$

$$\alpha_i = \frac{1}{\nu l} - \beta_i \leq \frac{1}{\nu l}, \qquad \sum_i \alpha_i = 1 \tag{8}$$

ラグランジュ関数 (6) に (7), (8) を代入し，(2) を使用すると，次の最適化問題を得る．

$$\min_{\boldsymbol{\alpha}} \frac{1}{2}\sum_{ij} \alpha_i\alpha_j k(\mathbf{x}_i,\mathbf{x}_j) \tag{9}$$

制約条件：

$$0 \leq \alpha_i \leq \frac{1}{\nu l}, \qquad \sum_i \alpha_i = 1 \tag{10}$$

3.2 SVDD

一クラスの分類問題で別の手法として使用される SVDD では，与えられた学習データを囲むなるべく小さな超球によりデータを分離する [Tax]．あらかじめ与えられた学習データの多くが超球の内部に含まれることを維持しながら，超球の中心 $\boldsymbol{a}$，半径 $r > 0$ について，超球の体積を最小化するものを求める．

これは，次の最適化問題で表される．

$$\min_{r,\boldsymbol{\xi},\mathbf{a}} r^2 + \frac{1}{\nu l}\sum_i \xi_i \tag{11}$$

制約条件：

$$\|\boldsymbol{\Phi}(\mathbf{x}_i) - \boldsymbol{a}\|^2 \leq r^2 + \xi_i, \qquad \xi_i \geq 0 \tag{12}$$

ここでのラグランジュ関数は次のように書ける．

$$\begin{aligned} L(r,\boldsymbol{a},\boldsymbol{\xi},\boldsymbol{\alpha},\boldsymbol{\beta}) = {} & r^2 + \frac{1}{\nu l}\sum_i \xi_i \\ & - \sum_i \alpha_i\{r^2 + \xi_i \\ & - (\|\boldsymbol{\Phi}(\mathbf{x}_i)\|^2 - 2\boldsymbol{a}\cdot\boldsymbol{\Phi}(\mathbf{x}_i) + \|\boldsymbol{a}\|^2)\} \\ & - \sum_i \beta_i\xi_i \end{aligned} \tag{13}$$

変数 $r, \boldsymbol{a}, \boldsymbol{\xi}$ について導関数を 0 とすると，次が成立する．

$$\boldsymbol{a} = \sum_i \alpha_i \boldsymbol{\Phi}(\mathbf{x}_i) \tag{14}$$

$$\alpha_i = \frac{1}{\nu l} - \beta_i \leq \frac{1}{\nu l}, \qquad \sum_i \alpha_i = 1 \tag{15}$$

ラグランジュ関数 (13) に (14), (15) を代入し，(2) を使用すると，次の最適化問題を得る．

$$\min_{\boldsymbol{\alpha}} \sum_{ij} \alpha_i\alpha_j k(\mathbf{x}_i,\mathbf{x}_j) - \sum_i \alpha_i k(\mathbf{x}_i,\mathbf{x}_i) \tag{16}$$

制約条件：

$$0 \leq \alpha_i \leq \frac{1}{\nu l}, \qquad \sum_i \alpha_i = 1 \tag{17}$$

これは次の識別関数 f に対応する．f について，$0 \leq \alpha_i \leq 1/\nu l$ において 0 となる．

$$f(\mathbf{x}) = \mathrm{sgn}\left(r^2 - \sum_{ij} \alpha_i \alpha_j k(\mathbf{x}_i, \mathbf{x}_j) + 2\sum_i \alpha_i k(\mathbf{x}_i, \mathbf{x}_i) - k(\mathbf{x}, \mathbf{x})\right) \quad (18)$$

Radial Basis カーネル関数や，本検討で使用した線形カーネル関数のように，$\mathbf{x}-\mathbf{y}$ のみに依存するカーネル関数 $k(\mathbf{x},\mathbf{y})$ については，$k(\mathbf{x},\mathbf{x})$ が定数となる．これは，SVDD における最適化問題 (16), (17) と，one class SVM における最適化問題 (9), (10) とが等価になることを意味する．

4. 評価

4.1 前提および評価方法

模擬的な監視制御システムとして，図 1 に示すシステムを用いた．制御箇所の監視制御中央装置により，被制御箇所にある フィールド機器の監視と制御を行う．通信方式は TCP (Transmission Control Protocol) を使用している．

フィールド機器の状態を示す監視情報は，監視制御遠隔端末装置を通し，監視制御中央装置に送信される．制御箇所の操作卓では，監視制御中央装置から監視情報を定期的に受信して表示する．各監視情報は，定期的にフィールド機器から監視制御中央装置に送信される他，状態変化時にも送信される．

フィールド機器の制御を行う際には，操作卓に表示されたフィールド機器を選択し，監視制御中央装置，監視制御遠隔端末装置を通して対象機器を選択状態にする．選択状態が確認できた後，再び操作卓から対象機器の制御指令を送信する．指令は監視制御中央装置，監視制御遠隔端末装置を通してフィールド機器に伝達され，制御が行われる．制御がエラーなく完了した際には確認応答が返信され，操作卓で確認することができる．制御の結果変更されたフィールド機器の状態は，監視情報として送られる．

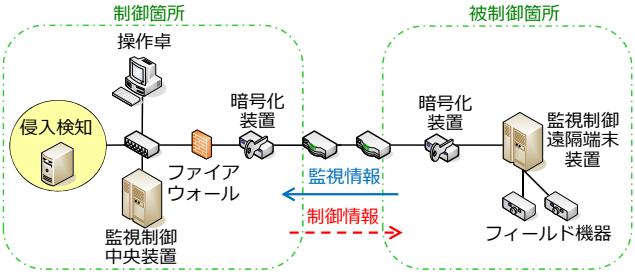

図 1 模擬的な監視制御システム

セキュリティ対策としては，制御箇所と被制御箇所間の通信の盗聴を防ぐ暗号化のために暗号化装置が設置されている．また，制御箇所への侵入を防ぐファイアウォールが，制御箇所の入り口に設置されている．侵入検知システムについては，監視や制御に伴う各種通信がもっとも多く観測できる箇所で，制御システム内のネットワークでの通信を監視できる位置に設置した．

フィールド機器から監視情報を監視制御中央装置に送信する通信，および操作卓からフィールド機器の制御を行い，その制御操作の確認を行う通信について，正規の通信を行った場合の通信パケットを取得し，侵入検知システムの学習データとした．学習用のデータ数は 1 万とした．

侵入検知精度を評価するための不正通信データは，実際のサイバー攻撃のデータが存在しないため，模擬的なサイバー攻撃により不正通信データを作成した．その際には，監視制御システムの通信ネットワークに攻撃者が接続できたものと仮定した．攻撃内容としては，正規の操作では考えづらい，連続的に制御を試みる操作や，監視情報や制御時の通信を通信経路の途中で乗っ取った上でその内容を変更した場合を想定した．情報通信システムで使用される侵入検知システムでは，通常はこのような攻撃を検知できず，監視制御システムで使用しているデータを解析し，それに応じた検知ルールを指定する必要がある．実際には，通信パターンの詳細な検討が必要となり，検知ルールのパラメータを適当な値に設定することは難しい．

通信パケット内のデータのうち，学習用データとして使用したものを表 1 に示した．各データは，監視制御システムの振る舞いに照らし合わせ，出現可能性が低いと考えられるものが 0 に近くなるように加工した．これにより，線形カーネルを使用する際に，出現可能性が低いパターンが原点に近くなる．

模擬的な監視制御システムから取得した正規の通信データを用いて one class SVM および SVDD の学習をそれぞれ行った．計算には，LIBSVM [Chang] および，PRTools [Duin] と

表 1 使用した学習データ

データ	説明
送信元アドレス(IP アドレス, MAC アドレス) 受信先アドレス(IP アドレス, MAC アドレス)	・システム内に存在する機器のアドレスは 1 ・システムと同一のアドレス空間は 0.5 ・システムと異なるアドレス空間は 0
送信元ポート番号 受信先ポート番号	・システム内で使用するポート番号は 1 ・システム内で使用する可能性のあるポート番号は 0.5 ・システムで使用しないと指定されているポート番号は 0
使用プロトコル 監視制御対象機器 ID 監視制御内容	・システム内で使用すると明示されている場合は 1 ・システムでの使用・不使用が明示されていない場合は 0.5 ・システムで使用しないと明示されている場合は 0
データ長	・システムで想定される範囲内の場合は 1 ・システムで想定される範囲に入らない場合は 0.5
通信パケット取得間隔	前の通信パケット取得時刻からの経過時間
監視制御対象機器使用後の経過時間	同一の対象機器を対象とした通信からの経過時間
同一監視制御内容実施後の経過時間	同一の監視制御内容を実施した通信からの経過時間
監視制御対象機器使用頻度	正規通信データセット内で対象機器が使用されている回数

DDtools [Tax] を使用した．νは比較的小さな値とした ($\nu = 0.01$)．学習の結果，外れ値として判定される通信内容の違いについて評価を行った．判定の内容によって，どのような通信が不正通信として判定され得るかについての知見を得ることが期待できる．

さらに模擬的な攻撃による不正通信のデータを利用し，one class SVM および SVDD のどちらの方式で学習した場合でも，正常通信および不正通信のそれぞれを正しく判定できるかを見た．

4.2 評価結果

正規の通信データを用いて one class SVM を学習させた結果，通信パケット取得間隔，監視制御対象機器使用後の経過時間，および同一監視制御内容実施後の経過時間が小さい通信が外れ値として判定されることを確認した．これは，今回対象とした監視制御システムにおいて，短い時間間隔で連続的に監視制御が行われることがないという想定と一致する．このような通信が不正に行われた際に検出できることが考えられる．

一方，SVDD については，one class SVM で判定された通信の一部と，監視制御対象機器使用頻度や監視制御対象機器使用後の経過時間が大きくなった通信についても外れ値として判定されることを確認した．この場合，一つの機器を連続的に多数回監視制御対象とした場合，あるいはほとんど使われない機器を監視制御対象とした場合に不正な通信として検出されることが考えられる．今回対象とした監視制御システムにおいて，一つの機器が連続的に多数回監視制御対象になることは想定されていないことと一致する．ほとんど使われていない機器については，実際にほとんど使われず，比較的長い時間間隔が空いた後に正規に監視情報が送信される場合と，不正に監視制御が行われる場合との両方の可能性が考えられ，正確な判別のためにはさらに別の入力が必要となる．

さらに模擬的な攻撃による不正通信のデータを利用した，正規通信および不正通信の判定については，one class SVM，SVDD のどちらの方式でも正しく判定することが確認できた．これは，想定した攻撃の通信において，監視制御対象機器使用後の経過時間および同一監視制御内容実施後の経過時間が特徴的になってしまったためと考えられる．通常の侵入検知システムでは，このような不正通信であっても，詳細な検討の上で設定を行う必要があるため，one class SVM および SVDD の適用は有効であると考えられる．

5. おわりに

模擬的な監視制御システムの通信データに one class SVM および SVDD (Support Vector Domain Description) を適用した．学習の結果，外れ値として判定される通信内容の違いについて評価を行った．どちらの方式でも，通信パケット取得間隔，監視制御対象機器使用後の経過時間，および同一監視制御内容実施後の経過時間が小さい通信が外れ値として判定される．SVDD では，監視制御対象機器使用頻度や監視制御対象機器使用後の経過時間が大きくなった通信についても外れ値として判定される．

また，模擬的なサイバー攻撃を行った不正通信データを用い，どちらの方式でも正常通信と不正通信を正しく判別できることを確認した．これらのことから，one class SVM および SVDD の，監視制御システムの侵入検知への適用は有効であると考えられる．

正常通信と不正通信の判別時には，監視制御対象機器使用後の経過時間および同一監視制御内容実施後の経過時間が影響していることから，今後はシーケンスを考慮した上での判別について検討していく．

参考文献

[Chang 2001] Chih-Chung Chang and Chih-Jen Lin: LIBSVM: A Library for Vector Machines, http://www.csie.ntu.edu.tw/~cjlin/libsvm, 2001.

[Chen 05] Wun-Hwa Chen, Sheng-Hsun Hsu, Hwang-Pin Shen: Application of SVM and ANN for intrusion detection, Computers & Operations Research 32, pp. 2617-2634, 2005.

[Duin 2007] R.P.W. Duin, P. Juszczak, P. Paclik, E. Pekalska, D. de Ridder, D.M.J. Tax, S. Verzakov: PRTools4.1, A Matlab Toolbox for Pattern Recognition, Delft University of Technology, 2007.

[Kiuchi 2009] M. Kiuchi, Y. Serizawa: Security Technologies, Usage and Guidelines in SCADA System Networks, ICCAS-SICE 2009, 2009.

[Kiuchi 2009] M. Kiuchi and Y. Serizawa: Customizing Control System Intrusion Detection at the Application Layer, SCADA Security Scientific Symposium 2009, Digital Bond Press, 2009.

[Onoda 2006] T. Onoda, H. Murata and S. Yamada: Non-Relevance Feedback Document Retrieval based on One Class SVM and SVDD, International Joint Conference on Neural Networks, 2006.

[Schölkopf 1999] B. Schölkopf, J. Platt, J. Shawe-Taylor, A. Smola, and R. Williamson: Estimating the support for a high-dimensional distribution, Microsoft Research, One Microsoft Way Redmond WA 98052, Tech. Rep. MSRTR-99-87, 1999.

[Tax 2004] D. Tax and R. Duin: Support vector data description, Machine Learning, vol. 54, pp. 45–66, 2004.

[Tax 2009] D.M.J. Tax: DDtools, the Data Description Toolbox for Matlab, http://homepage.tudelft.nl/n9d04/dd_tools.html, 2009.

[Zhang 2007] R. Zhang, S. Zhang, S. Muthuraman, J. Jiang: One Class Support Vector Machine for Anomaly Detection in the Communication Network Performance Data, 5th WSEAS Int. Conference on Applied Electromagnetics, Wireless and Optical Communications, 2007.