社団法人 情報処理学会 研究報告
IPSJ SIG Technical Report

2008－DPS－134 (13)
2008－CSEC－ 40 (13)
2008／3／6

エージェントシステムにおける不正エージェント排除のセキュリティー方式

安福 広† 小泉 寿男† 澤本 潤‡ 辻 秀一†††
†東京電機大学 理工学研究科 情報システム工学専攻
‡岩手県立大学 ソフトウェア情報学部 ソフトウェア情報学科
†††東海大学 情報理工学部 情報メディア学科

概要： インターネットを代表としたネットワークの普及に伴い、エージェント技術といった新技術が研究されている。一方、エージェントが持つ情報のセキュリティー保護が必要不可欠、という問題が持ち上がって来た。エージェント技術普及の為にはセキュリティーの研究が必要不可欠である。本研究では、マルチエージェントシステムの保護を狙いとし、なりすましに対抗するセキュリティー機能について検討した。結果、グルーピングや相互監視によってマルチエージェントシステム内の不正エージェントを排除する方法を提案する。本方式では、三層構造を用いてマルチエージェントシステムのセキュリティーを保護する。その三層とは IDS 層・認証層・不正エージェント排除層である。

A Security Method remove False Agent at Multi Agent System

Hiroshi Yasufuku† Hisao Koizumi† Jun Sawamoto‡ Hidekazu Tsuji†††
†Department of Computer and System Engineering
Graduate School of Science and Engineering, Tokyo Denki University
‡Faculty of Software and Information Science, Iwate Prefectural University
†††Department of Information Media Technology
School of Information Science and Technology, Tokai University

Abstract: **With the spread of networks like Internet, a new technology such as the agent technology is researched. On the other hand, a problem about the protection of the information security of an agent had arisen. A study of security is essential for the spread of the agent technology. In this paper, to protect the multi-agent system, we examine a security function against an impersonation attack. We propose a method to exclude malicious agents in the multi-agent system using grouping and mutual monitoring techniques. In this method, we consider three levels of structure to protect security of a multi-agent system. They are the IDS layer, authentication layer, and the malicious agent exclusion layer.**

1．はじめに

近年ネットワークが発達し端末同士の情報交換を行う環境が整ってきている。それに伴い、情報交換を用意にするための研究が発達してきている。また、情報のデータ化も進んできており、あらゆる情報がコンピュータに保存されるようになってきた。そのため情報の複写が容易に行えるようになっており、ネットワークを介しての情報の盗難や破壊が起こっている。攻撃手法として、例えばウィルス・トロイの木馬・DOS 攻撃・バッファオーバーフロー等が存在する。これらの危機から情報を守るためにセキュリティーの研究も進んできており、様々な対策技術が生み出されている。例えば IDS やハニーポット・フォレンジック等がそれにあたる[1]。

又、近年は情報交換技術としてｗｗｗやIRCを始め、P2P 等々、いろいろな情報交換技術が発達してきている。エージェントはユーザーの情報取得行動を代理して行えるソフトウェアであり、理想のエージェントはユーザーの意思を予測し、ユーザーが欲しがると思われる情報を手に入れるために自立して動作する。

しかし、エージェントシステムはユーザーのエージェント同士が通信を行いながら目的を達成するという意味でクライアントサーバーモデルを脱却しており、次世代のサービスを担うことになるプログラムであると言える。そのため、マルチエージェントシステム[4][5]のようなサービスが広まって行く可能性は十分に考えられる。しかし、セキュリティー対策は www 等や LAN 等のネットワークレベルでは進んできているが、エージェント分野ではセキュリティーの研究というのはまだまだ発展途上であり、エージェント普及の障害になっていると考えられる。

本研究では特にマルチエージェント環境におけるセキュリティー保護に焦点をあて、様々な攻撃手法に対抗できるセキュリティー機能の検討を行う。具体的には、防御段階を２層・３段階に分けて保護し、それぞれの段階をIDS層・認証層・不正エージェント検知層とする。特にこの中でも、不正エージェント検知層による「なりすまし」エージェントの検出をメインに扱う。そして、マルチエージェントネットワークに侵入した、不正エージェントを排除できる仕組みの構築を目指す。

２．エージェントとセキュリティー

現在、セキュリティーを脅かす攻撃手法は、多々存在している。たとえば、「トロイの木馬」「ウィルス」といったプログラムが悪さを行う形のものから、「なりすまし」「クロスサイトスクリプティング」「バッファオーバーフロー」等々のネットワークを介したハッキング攻撃まで様々である。これらの攻撃を防ぐために、FW やIDS・ハニーポット・フォレンジック・ウィルス対策ソフト等が開発され、対策にあたってきている。

エージェントは、ユーザーの代理人として自立的に判断して動作するソフトウェアの事である。交渉や情報収集をユーザーに代理して行う事が出来、ユーザーの操作を予測して、支援を行う事も出来るソフトウェアである。図１にエージェントの連携概略図を示す[4][5]。

エージェントはその動作形態に応じて数種類に分類されている。代表的な動作形態として、モバイルエージェントとマルチエージェントが存在する。モバイルエージェントは端末間を移動しながら動作し、単体で目的を達成する。マルチエージェントは無数のエージェントが、連絡を密に取り合う事で連携し、システム全体で一つの目的を達成するエージェントシステムの事であり、協力を行う事で目的を達成する協力型と、競争する事で達成する競争型が存在する。

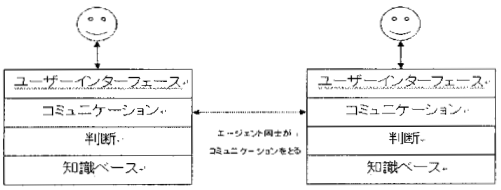

図１ エージェント連携概略図

これらのエージェントの中には、エージェントネットワークの機能停止や、作業効率の低下、情報の盗難といった不正行為を目的として、エージェントネットワークに侵入している「なりすましエージェント」が存在する。例えば、不正に入手したIDとパスワードでの侵入や、正規エージェントを改変し無能なエージェントにする事等でシステムに不正なエージェントが侵入した状態にできる。本研究では特にマルチエージェントシステムにおける「なりすまし」対策について取り扱い、なりすました「不正エージェント」の排除を対象とする。

３．不正エージェント排除のセキュリティー方式

3.1 概要

エージェントのなりすましを防ぐにあたって、本研究ではなりすましのレベルに的確に対処するために防御手法を複数段階用意し、より的確になりすましに対抗する事を目指す。複数用意するのは、ネットワークに侵入する手段や対象は複数存在するからである。LAN ネットワークに対する「なりすまし」とエージェントネットワークに対する「なりすまし」では対処方法が違う[7]。又、ネットワークに侵入する方法も多数存在する。IDとパスワードを使って正門から侵入する方法や、裏口から侵入する方法など、侵入方法も色々である。

これらに的確に対処するために、本研究では侵入を行う対象毎に防御手法を分割して２層構造をとする。そのうちの一層目では、ＯＳＩ参照モデルにおける物理層からインターネット層に相当する部分のなりすまし対策を行う。二層目では、「なりすまし」対策を行う事とする。特にこの層は、エージェントネットワークに侵入を防ぐ構造と、エージェントネットワークシステムへ侵入してしまった場合の検出・正規エージェントが変質し不正エージェントになってしまった場合の検出を行う構造の二つに分割する。図２に機能分割の概略図を示す。

エージェントセキュリティー層	不正エージェント検知層	侵入・変質した不正エージェントを防ぐ
	認証層	エージェントネットワークへの侵入を防ぐ
IDS 層		LAN ネットワークへの侵入を防ぐ

図２ 機能分割図

3.2 IDS 層の機能と事例

IDS 層は一般的な物理ネットワークへの侵入を防ぐ層である。この層の機能によって、物理ネットワークへの侵入は不可能になる。この層が存在によって、LANに端末を接続することによって、ネットワーク上に流れているパスワードやIDを盗む事が出来なくなる。この層が提供する機能は、登録されていない端末をネットワークに接続させない事である。

実現方法として、ARP と DHCP を監視し、IP アドレスとMACアドレスの対応表を作る事によって実現して

いる。DHCP は、コンピュータがネットワークに接続するための情報を取得するプロトコルであり、ブロードキャストで通信が行われる。そのため、ネットワークを流れるパケットを監視する事によって、IP アドレスと MAC アドレスの対応表を作る事が出来る。ARP は、IP アドレスから MAC アドレスを取得するためのプロトコルであり、ブロードキャストで通信するプロトコルである。ARP の監視結果と DHCP を監視して取得した IP-MAC 対応表と照らし合わせることで不正な端末を突き止められる。突き止めた不正端末はネットワークから切断する。本研究ではこの層の実現方法として、IPS ソフトを利用する事を考え、Net Skate Koban[6]という IPS ソフトを評価した。IPS とは IDS では検出機能だけであり検出するだけであるが、IPS ソフトは検出と対策を行う事が出来るソフトウェアである。

3.3 認証層の機能

認証層はエージェントネットワークへの侵入を防ぐために存在する層である。この層は、実際には目には見えず、現実世界には存在しない論理的なネットワーク（マルチエージェントネットワーク）への侵入を防ぐために存在する。この層の機能によって、この論理ネットワークへの侵入を限りなく難しくできる。

本方式では、マルチエージェントシステムへの接続時に認証サーバーを通してしか参加できないようにする事によって、マルチエージェントシステムに異分子が入り込まないようにする。これによりシステムに参加しているエージェントは全て認証サーバーによって、管理され、認証サーバーに問い合わせればそのエージェントが正規の物であるかどうかが分かる仕組みにした。これによって、正規でないエージェントはマルチエージェントシステムへの侵入がかなり難しくなる。しかし、アカウントが盗まれてしまうと進入されてしまう弱点があるため、アカウントが盗まれないように認証通信を強固な暗号で行う事とした。

エージェント認証サーバーは次のような通信を提供し、エージェントの管理を行う。

- 通信は全て暗号で行う。これによりパケット監視による情報の盗難・不正侵入を難しくする。
- 暗号は定期的に変更する。これによって暗号の解析に対抗する。
- 認証サーバーにエージェントを認証させるログインは暗号を用いて行う。
- 必要ならば端末の制御を行い、特定の IP 内からしかアクセスできないようにする。

本方式における安全な通信路でエージェントと接続を確立する方法の概要を図３に示す。

① マルチエージェントシステムに接続したいエージェントは認証サーバーに接続する。
② 認証サーバーは公開鍵をエージェントに送信する。
③ エージェントは、ランダムに暗号を作成し、サーバーの公開鍵を使って作成した鍵を暗号化する。
④ 認証サーバーは、エージェントから送信されてきた鍵を使ってエージェントとの通信を開始する[3]。

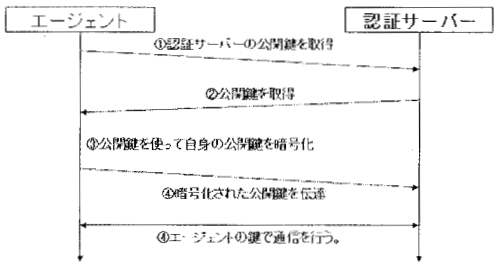

図３ 通信路確保

次に通信路が確保された後の認証手順を図４に示す。認証サーバーは、ID 情報認証情報を要求し照合を行い認証する。認証の結果正規エージェントと判断された場合は、その時点においてエージェントネットワーク内で使われている通信用共通鍵を送信する。送信された鍵を使って他のエージェントとの通信を開始する。通信用共通鍵は一定時間ごとに変更する。それによって鍵が解読された場合や、漏洩した場合に備える

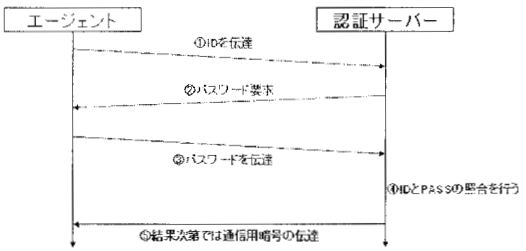

図４ 認証通信内容

また認証サーバーには IDS 層や不正エージェント検知層との連携機能を設けてある。IDS 層との連携では、不正エージェントが動作していた端末を調査・除外や不正端末からの接続者が居ないかを互いに連絡しあう。又、不正エージェント検知層との連携では、検知したエージェントにテストを行い、最終的な不正エージェントであるが否かの判断は認証サーバーが行う。

3.4 不正エージェント検知層の機能

3.4.1 不正エージェント検知

不正エージェント検知層は予期しづらい方法等でエージェントネットワークに侵入したなりすましエージェントの検出や、正規エージェントがウィルス等の要因

によって変質し、不正なりすましエージェントになった場合を検出するための機能を持つのがこの層である。この層の存在によって、成りすましているエージェントを見分けネットワークの安全性を保てる。ネットワーク層の盗聴防止策・なりすまし防御策を破られてしまった場合や、なんらかの方法でネットワークに参加されてしまった場合に、すばやく「なりすまし」エージェントを見付ける事によって被害少なくすることが出来る。いわば、人間社会における警察のような役目を持っている層である。

本方式では、各エージェント・認証サーバーは不正なエージェントを検出するために以下の機能を持たせる。これらの機能は、エージェント同士が互いに監視する形になっており、これにより不審エージェントを見つける事が可能になる。

① 全エージェントは動作を行うたびに、必ず認証サーバーに報告を行う。認証サーバーは報告を元に、各エージェントの動作ログを作成する。エージェント同士が通信を行った場合は必ず二つのエージェントにログが残る。そのため、片方が報告を怠るとログに相違が生じる。このログでは通信内容やエージェントの内部データまでは報告しない。報告は通信や目的等の動作報告に限る。

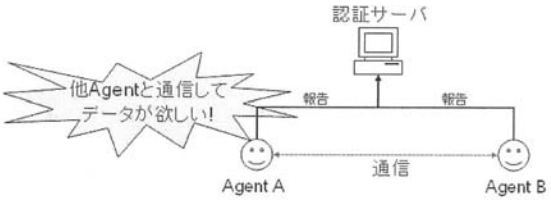

図5 ログ保存概略

② 各エージェントは相手が自分にとって有効かを判断する能力を持ち、全ての通信相手に対して友好度を設定する。情報を聞くだけで、自分からは情報をもたらさない等の怪しい動作をした場合や、相手エージェントとの通信が自分によって有益でないと判断した場合には、この友好度が下がる。友好度が一定以下になった場合には接続を切断し認証サーバーに報告を行う。認証サーバーは、報告が一定数たまったところでテストを行い、不正エージェントかどうかテストを行う。図6はエージェントネットワークの模式図を表している。この例では、真ん中の赤いエージェントとの通信を切断した右側の二つのエージェントが認証サーバーに報告を行っている。認証サーバーは報告を受けたエージェントを不正エージェント候補として記憶しておく。

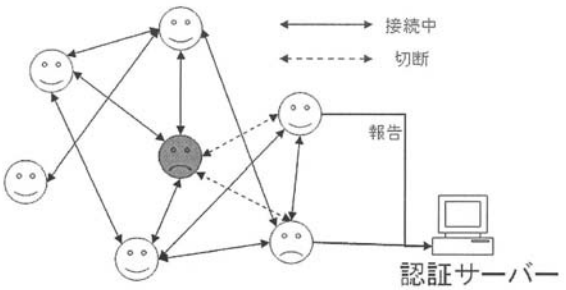

図6 相互監視

③ エージェントに良く一緒に作業するエージェントや、作業の方向性が似ているエージェント同士でグループを作成する機能を搭載し、エージェントの作業や目的毎にグループを作成する。相互監視方式では、妨害動作をせずに情報の盗難だけを目的にした不正エージェント等を検出できない。しかし、こうした表面上大人しいエージェントは普通のエージェントとは目的に相違があるため、目的ごとにグルーピングを行うと、不正エージェントだけ浮き出た存在になると考えられる。また、グルーピングを行う事によって、通信が密になり作業効率の上昇が考えられる。又、作業に必要な情報交換がグループの中でのみ密に行われることによって、盗難をたくらむ不正エージェントに情報が行き渡らないため、情報盗難の危険性が低下すると考えられる[2]。図7はグルーピングされた様子を表している。この図では、真ん中の赤いエージェントが不正エージェントをあらわしている。どのグループにも情報を盗みに行こうとしているため、どのグループにも所属できない。

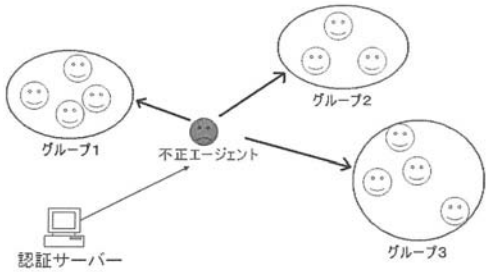

図7 グルーピング機能

3.4.2 不正疑惑エージェント検出のためのテスト方式

これら三個の方法によって疑わしい不正エージェント候補をピックアップする。不正エージェントの嫌疑がかけられたエージェントは認証サーバーによってテストされ、最終的に不正エージェントであるかどうかを判断される。テスト方式としては以下の３個を考案した。

① 作業応答方式

協力型マルチエージェントシステムの場合は擬似的なシステムを用意し、対象エージェントをそのシステム内部で共同作業させ、エージェントの作業貢献度を測り判断を行う。不正エージェントであった場合で、特にシステム妨害が目的の不正エージェントの場合、協力を行わないと考えられるので、判別できると考えている。競争型エージェントの場合は、正常に競争が行われているかで判断を行う。

② おとり捜査方式

エージェントの今までの通信ログから「もし不正エージェントであるならば、何を行いたいのか」を推論しテストする。不正エージェントにも何かしらの目的が存在するはずである。例えば、情報の盗難や、システムの機能停止、システムにバックドアを作る等である。その推論を元に不正エージェントを隔離された擬似ネットワーク環境と通信を行わせ、怪しい挙動をしていないか調査する。又、ネットワークに不正エージェントが欲しがるような情報をまき、情報やエージェントが飛びつく頻度を見て判断を行う

③ IDS ログ対比調査方式

認証サーバー内に保管されているログを詳細に調査し、嫌疑をかけられているエージェントの動作に不審な点がないか調査する。例えば、動作の報告漏れがないか検証する。又、IDS 層でのログを参照し、嫌疑がかかっているエージェントの端末が外部ネットワークと怪しい通信をしていないか、パケットに不審な部分は無いか調査する[8]。

この三つのテスト方式によって不正であるか否かを判断する。これらによって不正なエージェントであると判断された不正エージェントはネットワークから切断する。また、不正なエージェントが認証に使った ID とパスワードを無効にし、再ログインを防止する。また、不正エージェントが発見された場合はエージェントネットワーク全体で使われている通信暗号を変更し、ネットワークの内部への再接続を防ぎ、パケットを盗み見られても情報が盗難されないようにする。

4．構築

実装は各層によって実装方法に異なる手法を使用している。それは層によって求められている性能・機能・実装対象・動作環境が違うからである。

IDS 層は IDS･IPS ソフトウェアを実現方法として評価した。具体的には、Net Skate Koban を実現方法として使用した。Net Skate Koban の概略図を図８に示す。Net Skate Koban は Windows で動作させた。

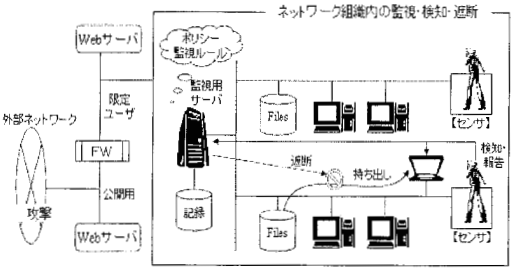

図８ Net Skate Koban 概略図

Net Skate Koban はネットワークの監視や管理・不正端末の切断を行う事ができる。また、その接続ログを詳細に保存でき、ログエージェントを使うことで、詳細なログ解析を行う事ができる。また、ネットワーク状態の検証レポートを出す事ができ、認証層・不正エージェント検知層との連携も問題なく行えると判断できた。そのため、この製品を使うことで、IDS 層で想定していた機能はほとんど実装できる事が確認できた。そのため、IDS 層は Net Skate Koban を利用して実装を行った。

認証層・不正エージェント検知層においては、認証サーバー部分は C++での実装を行った。これは、沢山のエージェントを管理する必要があり仕事量が多いため、処理速度・応答速度を要求されるからである。暗号部分の実装は、今回はライブラリを用いて実装を行った。エージェント同士の通信暗号には BlowFish を用いた。公開鍵には PGP を用いた。開発は Windows で行い、コンパイラは Visual Studio を用いた。エージェント実装部分は、最終的にはシステム内の全てのエージェントに組み込む必要があるため、システム内に存在する様々なプラットホームで動作しなくてはいけない。そこで、マルチプラットフォームで動作する JAVA で実装を行う。後から組み込む事が出来るように、ライブラリ形式で実装を行う事を考えている。開発は Windows で行った。

各エージェント同士の通信とエージェントと認証サーバーの通信は TCP を用いて実装を行う。

5．評価

評価は研究室のネットワークを用いて行っている。研究室の PC 上でエージェントと認証サーバーを実行し、エージェント数２０の競争型マルチエージェントネットワークを作成した。２０とした理由は５～６個のエージェントグループ３個と不正エージェント２～５個を作成できる最小構成だからである。グループ化進めるために、エージェントは同時に５～８台のエージェントとのみ通信し、全てのエージェントと同時に通信を行わない。最初の接続相手は認証サーバーがランダムに選ぶ事

にした。エージェントの目的はユーザーの作業支援とした。この条件下のエージェントネットワーク上で次の４点について実験を行い、提案方式の能力を評価している。

① 認証堅強度評価

エージェントがエージェントネットワークへのログインする際に認証サーバーとやり取りするパケットから、ＩＤ・パスワードが推測できるかを評価する。暗号化されたパケットを取り出し、解析にかかる予想時間を求める。予想時間が十分に長ければ強度として問題ないと判断する。また、エージェント認証に最適な暗号系を模索する。（認証層の評価）

② 盗聴による情報盗難耐性評価

エージェントネットワークを流れている情報を溜め込み、盗む事を目的としたエージェントを意図的に作成し、そのエージェントを不正エージェント検知層の機能で判別できるかを評価する。次の２種類のエージェントを判別できるかを判断基準とする。（不正エージェント検知層の評価）

- 情報を盗むだけでまったく作業をしないエージェント
- 共同作業にも参加するが、情報を盗み外部に情報を持ち出しているエージェント

③ 作業妨害耐性評価

エージェントネットワークの効率を下げるために存在するエージェントへの耐性を評価する。協力作業を行わず、他の作業を邪魔するエージェントを作成し、そのエージェントを不正と判別を受けるか否かを評価する。これについても２種類の不正エージェントについて判断を行う。図９は赤で表した、作業を妨害するエージェントを右上の認証サーバーが不正嫌疑をかけている。（不正エージェント検知層の評価）

- 共同作業への参加を持ちかけ、実際には作業を行わなかったり、虚偽の作業結果を報告したりするエージェント
- マルチエージェントネットワークに大量に侵入しネットワーク全体での正常エージェントの数を減らす。それによりシステム全体での作業効率を下げる攻撃に対する耐性

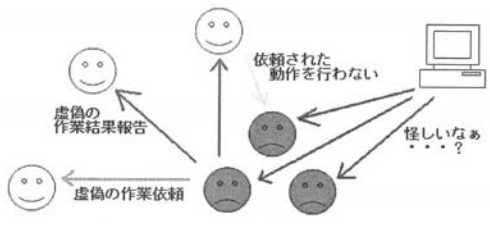

図９ 作業妨害耐性評価概略図

④ 不正エージェント検証テストの評価

不正エージェントと正常エージェントに、不正エージェントの嫌疑をかけ、認証サーバーにテストにさせる。そのテストの結果、正常と不正を正しく判別し、仕分ける事が出来るかを評価し、提案したテスト方式で不正・正常が判別できるか判断する。（不正エージェント検知層が行う不正判定テストの妥当性評価）

これらの実験により、各エージェントや認証サーバーの対策モジュールが正常に動作しているかを評価する。

６．まとめ

マルチエージェントシステムを守るシステムの枠組みを考案した。具体的には、IDS 層・認証層・不正エージェント検知層の３個の仕組みを設計した。IDS 層は、Net Skate Koban の評価を行い IDS 層の機能として問題ないことを確認し、IDS 層として実装した。認証層の認証プロトコルの設計も行い、パスワードが簡単に盗まれないプロトコルを構築した。また、内部に侵入した不正エージェントを焙り出す方法を考案し、正規エージェントと不正エージェントを見分けるテストの方法を提案した。今後構築を行い、考案方法で不正なエージェントの防御・排除を行う事が出来るのか評価を進めている。

参考文献

[1] 小松　文子：「プライバシ保護のためのアーキテクチャ」，情報処理,Vol.48,No.7

[2] 原　章・長尾　智晴：「自動グループ構成手法 ADG によるマルチエージェントの行動制御」，情報処理学会論文誌，Vol.41,No.4

[3] 鈴木　秀一：「秘密鍵を特定できない暗号」コンピューターセキュリティー学会,2000 年 7 月 25 日

[4] 西田　豊明・木下　哲男・北村　康彦・間瀬　健二：オーム社，「エージェント工学」，2002 年 7 月 15 日

[5] 岡田　謙一・西田　正吾・葛岡　英明・仲谷　美江・塩沢　秀和：オーム社，「ヒューマンコンピュータインタラクション」，2002 年 8 月 15 日

[6] Cyber Solution Inc ，「Net Skate Koban v3.0 ユーザマニュアル」「リファレンスマニュアル」一式，2006 年（http://www.cysol.co.jp/products/netskatekoban/index_j.html）

[7] 寺田　真敏・甲斐　賢・熊谷　仁志「不正な TCP コネクション確立に関する一考察」，コンピューターセキュリティー学会,1999 年 7 月 9 日

[8] 竹森　敬祐・田中　俊昭・清水　晋作・中尾　康二，「不正侵入者に検知される事なくおとりデータ領域へと誘導するおとりシステムの実装評価」，コンピューターセキュリティー学会，2001 年 2 月 21 日