Applying Emergent Immunological Concepts to Network Intrusion Detection

Mario J. Antunes[1], Manuel E. Correia[2]

[1] Polytechnic Institute of Leiria - School of Technology and Management,
Morro do Lena - Alto do Vieiro
2411-901 Leiria, Portugal

mario.antunes@estg.ipleiria.pt

[2] Computer Science Department - Faculty of Science - University of Porto,
Rua do Campo Alegre, 823
4150-180 Porto, Portugal

mcc@dc.fc.up.pt

Abstract

This paper outlines the use of theoretical immunological concepts in the deployment of a novel Network Intrusion Detection Systems (NIDS) framework. It describes a research project that aims at defining a new extension to the Common Intrusion Detection Framework (CIDF). During the last decade several immunological concepts have been used in intrusion detection systems through Artificial Immune Systems (AIS) implementations, such as clonal selection, negative selection and self-non-self distinction. It has been recognized that these systems have some limitations, therefore there is an actual need to develop new and more robust systems incorporating the more recent developments in theoretical immunology, being the Danger Theory (DT) one of the most promising. In this paper we begin by presenting some important principles and concepts we think are most relevant to the description and categorisation of intrusion detection systems (IDS). We then proceed to describe the main benefits that can be obtained from an artificial immune system approach for IDS, stressing the new trend based on danger theory. We conclude by presenting a novel extension to the common intrusion detection framework (CIDF), stressing some of the main benefits that can be obtained by using an immunity-inspired approach based on Danger Theory.

1 Introduction

A Network Intrusion Detection System (NIDS) main activity consists in analysing the flow of packets in the network and identify which ones are part of an attack. One of the major problems related with NIDS deployment is the detection of new kind of attacks that have not occurred previously. There are several well documented approaches to detect and try to learn from new forms of attack, mainly statistically based [17]. Unfortunately none seems to provide a completely satisfactory answer. In this article we propose an approach based on biologically inspired concepts and algorithms, namely the ones related with the human immune system [7]. We take advantage of concepts, ideas and algorithms

based on the theoretical biological models of the human immune system (AIS - Artificial Immune Systems [14, 13]), and apply them to intrusion and anomaly detection on computer networks.

The previous and well accepted work on intrusion Detection Systems (IDS) inspired us to define an extension to the Common Intrusion Detection Framework (CIDF) [22, 8], by refining and enhancing its capabilities to work in an adaptive and self-adjusting manner. We recently proposed the Immunity-Inspired Intrusion Detection System (I3DF)[4]. It is still a work in progress research project based on the application of immunological algorithms and methods to traffic flows classified as normal and through the definition of a normality profile, based on the relationships between hosts in the network (clustering approach)[27]. In this paper we concentrate on the biologicall inspired aspects of this project and leave the details on traffic flow classification to other publications[4].

In section 2 we revise some basic concepts of intrusion detection systems and in section 3 we introduce and explain some of the fundamentals behind biological immune systems. We then proceed to section 4, where we present artificial immune system models that have been successfully applied to intrusion detection systems. In Section 5 we explain the actual developments done so far in Danger Theory (DT)[2] and describe in some detail our framework, I3DF (Immunity-Inspired Intrusion Detection Framework). In this section we also present the CIDF and introduce a new proposal for the use of the DT in the scope of intrusion detection. Finally, in section 6, we detail some conclusions, reflecting the study we have done so far and discuss some directions for future development and research.

2 Intrusion Detection Systems

Intrusion can be seen as a set of actions that attempt to compromise a secure property. Intrusion detection is the process of monitoring relevant events that occur in a computer-based information system. The main goal of intrusion detection is thus to positively identify all occurrences of actual attacks and, at the same time, to not be mistaken by regular events and distracted by the signalling of false attacks [29]. The intrusion detection system's main goal is to detect unauthorised use, misuse and abuse of computer systems by both system insiders and external intruders.

There are several ways to identify and categorise existing IDS. If we begin by considering the source from where an IDS gets its information, these systems can be classified as Network IDS (NIDS), Host IDS (HIDS) and Hybrids.

Broadly speaking, there are basically two approaches for the manner in which an IDS identifies potential intrusions: *anomaly detection* and *miuse detection* [16].

Anomaly (behaviour-based) detection bases its decisions on a profile of normal network or system behaviour. It starts by building a model for normal system behaviour. This is denoted by what is called the normal activity profile.

It then proceeds by looking for anomalous activities, which by definition are activities that do not match the previously established profile. An intrusion is thus a deviation from the normal activity profile. These anomaly detection systems make effective use statistical analysis, predictive pattern generation, neural networks and genetic algorithms[5].

The misuse detection (knowledge-based) based systems examines network and system activity, comparing the data collected by the IDS with the contents of a database, looking for known misuses. The database contains the signatures of known attacks in the form of rules. If a match is found, an alert is generated and all the events that do not match any signature are considered not intrusive. These systems are based on the use of expert system technology, state-transition analysis and pattern matching algorithms [5].

Both of these methods have strengths and weaknesses. In one hand, misuse-based systems generally have a very low rate of false positives but cannot identify novel attacks, leading to high false negative rates. On the other hand, anomaly-based systems are able to detect novel attacks but currently produce a large number of both false positives and false negatives[5]. This problems are due to the inability of current anomaly-based techniques to deal adequately with continuous changes in network environments. This is a clear indication for the need to find and apply new paradigms that can better cope with legitimate changes in computer networks and systems usage over time, meaning that any kind of profile for normal behaviour also needs to be dynamic in nature.

The application of biological immune system concepts and algorithms provides the system with the innate capability to learn and memorise past events. This increases the quality and resilience of an IDS by providing it with the ability to react to new and never encountered attacks.

3 Biological Immune System

The biological immune system is a very complex multi-layered structure that evolved to protect and defend the body from microorganisms (pathogens) that can cause diseases, such as virus and bacteria [7]. Antigens are substances (usually proteins) identified as foreign by the immune system. They stimulate the release of antibodies to destroy the pathogens and are composed by a set of cellular components that interact with each other to react against an intruder (Figure 1).

The immune system possesses two main levels of defence: *innate* and *adaptive*. The *innate* level of defence is a direct result of each person's individual genetic information. It has no learning mechanisms and always reacts in the same way to intruders. The *adaptive* immune level of defence, on the other hand, recognises antigens according to the previously acquired memory of past intrusions and reacts adaptively to new or similar events. In the adaptive system, the *specificity* refers to the binding process of an antigen by a cell. Each cell has a receptor that can only recognise one specific antigen. These cells are the *leukocytes* (white blood cells) and correspond to the main cellular components

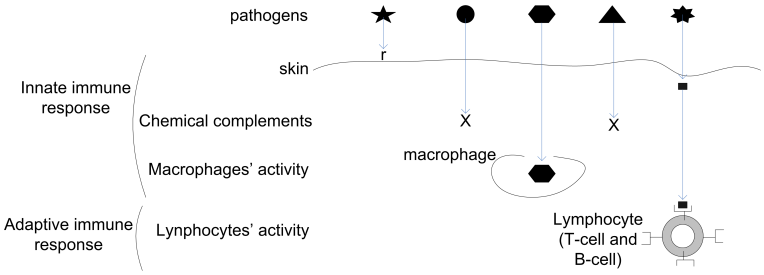

Figure 1: Multi layered structure of immune system

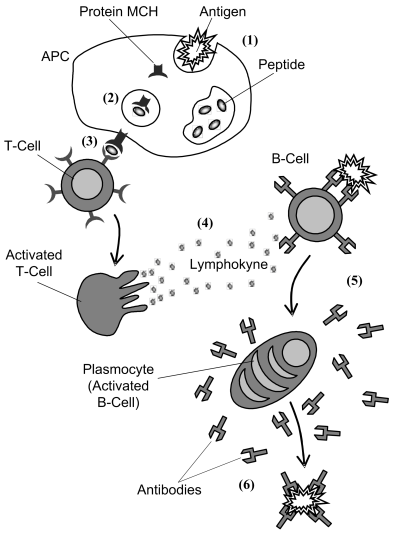

Figure 2: The immune system in action.

of the immune system. Macrophages are the main cells in the innate system destroying the antigens they are able find and bind with within the body. On the other hand, B-cells and T-cells are responsible for the adaptive immune system behaviour and activity. The immune system acts as a whole in the destruction of microorganisms. Figure 2 [14] illustrates its behaviour in the normal activity of detenction and reaction to invaders.

In (1) the macrophages, a specific kind of cells denominated APC (*Antigen Presenting Cells*) look for foreign antigens, destroying and fragmenting them in antigenic peptides. Some of these peptides bind to special proteins (2), called MHC (*Major Histocompatibility Complex*), being presented to the cell surface as a pair "MHC/peptide". In (3) the T-cells bind to this pair being stimulated and activated. After its activation the T-cells will activate also the B-cells (4), through the release of some chemicals (e.g. lymphokine). In (5) B-cells will be divided and differentiated in plasmocytes that will produce a high rate of

antibodies and *memory cells*. The antibodies produced will have the shape of the B-cell receptors, thus binding the antigen and neutralising the intruder (6).

The adaptive immune system has a wide variety of interesting characteristics that can be used in sophisticated computer applications, such as highly adaptive intrusion detection.

In immunological models the concept of *diversity* is related to the large variety of cells receptors that can possibly bind to all types of antigens or intruders. In computer networks, a potential attack would be identified by a specific receptor that would react through adequate security countermeasures. The system must therefore possess a large diversity of detectors to cope with as much attacks as possible. The *negative selection* process [24]allows the immune system to differentiate between self and non-self cells, thus preventing the body from start being attacked by its own immune system cells, which would lead to an autoimmune disease. An IDS can distinguish normal from abnormal activity using a similar method. For example, an open TCP port can be seen as a normal characteristic of the system by accepting connections in that port. All other connections to TCP ports not defined in the open state should be considered as potential intrusions.

Finally, the *theory of clonal selection* [14]explains why in each individual, antibodies are only produced for the antigens to which he has been previously exposed. Considering that a cell can only bind to a specific antigen, it is necessary to have a huge number of different cells to be able to efficiently cope with an infection. Thus, after being stimulated by an antigen, the cell is cloned into a multitude of copies, creating a huge group of cells capable of attacking that specific antigen. For example, in the case of B-cells, the clonal selection process produces two kinds of cells: the plasmocytes and memory B-cells. The former will be responsible for a large scale production of antibodies to fight the antigens. The later are memory cells that can remain in the body for long life periods, with the main function of reacting faster to a second similar attack.

In a network there are various sources of events that can be used in the evaluation of potential attacks. For example, various IDS sensors (cells) could be installed in the network and the hosts events (logs) should be collected and analysed. The IDS sensors can thus be seen as cells that work together in a distributed way, collecting and analysing events for further reaction. Finally, the "memory cells" must be generated by the IDS to guarantee a future faster reaction to similar attacks.

These are the most important immunological concepts deployed in artificial immune systems (section 4), with relevance to network intrusion detection. Nevertheless their application is definitely not an answer to all the problems faced today by intrusion detection systems. For example, the negative selection phase shows scaling problems when it is applied to real network traffic [23]. The computational complexity in an IDS grows exponentially with the number of systems (services) that need to be protected. It is thus necessary to find a diversified set of "receptors" that can provide adequate coverage and at the same time be computationally efficient. The mapping of the entire self or non-self receptor universe is an inefficient task, since they change over time and only a

small minority of non-self is harmful, whilst some self may cause damage [1]. To make matters worse, the self and non-self definitions are often ambiguous and not always applied in the best possible way[23].

4 Artificial Immune Systems applied to IDS

Network computer security can be seen as one of the most intuitive and popular fields where we can effectively use the biological immune system as a computing metaphor. In his seminal work, Forrest *et al.*[18] managed to take full advantage of some important characteristics of the immune system, such as diversity, adaptability, anomaly detection, multiple layers and identity by behaviour, among others, to engineer LISYS [6, 21]. This was one of the first successful network intrusion detection systems based on AIS. In [19] Forrest proposed a first approach to deploy AIS in network security, where the non-self is characterised as "undesired network connections". In [23] Kim identified three fundamental design goals requirements for network based intrusion detection systems: distribution, self-organisation and lightweight operation. He also concludes a typical AIS framework must include negative selection and clonal selection mechanisms and should take advantage of gene library evolution algorithms. He also presents an AIS incorporating the requirements and characteristics listed above, describes the developed architecture and shows some promising results of its application in a real local area network.

Dasgupta [10] proposed an agent-based framework for intrusion/anomaly detection and reaction in networked computers. The mobile agents are able to interact with each other by travelling around the network nodes and monitoring several parameters, such as the type of user and its privileges, amount of free memory and connection types. Other Dasgupta's contributions in computer security can be found in [11, 9].

In [13] De Castro defines the concept of Immune Engineering and proposes a general framework and a set of tools to be used in a wide range of applications. In [15] de Paula *et al.* proposed a prototype called (ADENOIDS) inspired by immune systems and featuring automated intrusion recovery and the automatic extraction of a signature for remote buffer overflow attacks.

In [3] Aickelin *et al.* presents a very complete survey of intrusion detection systems based on AIS developed thus far, stressing their weaknesses and defending the need to adopt a new paradigm, the Danger Theory. This derived theory is introduced in section 5.

Finally, in [12] it is possible to find an extensive and actual bibliography of the related work developed so far in the scope of artificial immune systems.

5 Danger theory

Although very recent, uncompleted and currently still surrounded by some controversy [28], the DT [25] is gaining increased popularity amongst theoretical

immunologists. The central point of the immune system is its ability to respond to foreign antigens and to not react to self molecules. In order to undertake this role the immune system needs to be able to differentiate between non-self, and possibly invaders, from self molecules. It is currently well established in classical theoretical immunology that the immune response is triggered when the body encounters something that is non-self or foreign, in a discrimination process known as self-non-self recognition [7].

There are however some natural phenomenons that cannot be completely explained by classical immunological theories. For example, there is no immune reaction to foreign bacteria in the food we eat although they are foreign entities. The successful transplants of foreign organs are also a good example of no attacks against foreign (non-self) tissues.

Besides the theoretical immunologists assumptions that the self-non-self distinction is made through the elimination of cells that react to the self, in a self elimination process [7], Matzinger's Danger Theory [25, 26] proposes that there must also be some kind of discrimination process that goes beyond that distinction. Thus, the immune system does not react to non-self but to *danger*, i.e, the 'foreignness" of the invaders is not so important for the immune recognition as the relative "danger" of these invaders.

The danger theory central idea is that the immune system does not react to non-self but to danger. The system discriminates "some" self and "some" non-self, which starts to explain why it is possible to cope with "non-self but harmless" and with "self but harmful" system aggressors [2] (for example, a tumour). The theory states that foreign cells (invaders) will induce the generation of specific cellular molecules (danger signals), by initiating cellular stress (cell death) in some unnatural way[25]. These molecules will trigger the immune response by being recognised by the APC (see Section 3). These signals encourage the macrophages to capture antigens in the neighbourhood and establish a "danger zone" around the alarm signal emitted by the distressed cell. Only those B cells producing antibodies that match antigens within the danger zone get stimulated and start the clonal expansion process. Thus, this new theory suggests that the immune system reaction to threats is based on the correlation of various signals reported by the immune system "sensors", readily providing a method of linking the threat directly to the attacker.

In an IDS context the danger alerts can be reported by the various sensors distributed within the network. Having received strong indications of a possible intrusion scenario, the AIS should send signals to other sensors in the neighbourhood, allowing a triggered action to the intrusion. These signals can be of two types: *apoptotic* and *nectotic*. The former corresponds to a normal death of a cell and in an IDS this would correspond to legitimate actions or some prerequisites for an attack. The later is an unregulated cell death process and in the context of an IDS it would correspond to actual damage caused by a successful attack [1].

Aickelin [1, 20] aims to investigate the correlation described above and transpose the danger theory to the realm of computer security. In his approach the self-non-self discrimination is still used but no longer essential, since the reac-

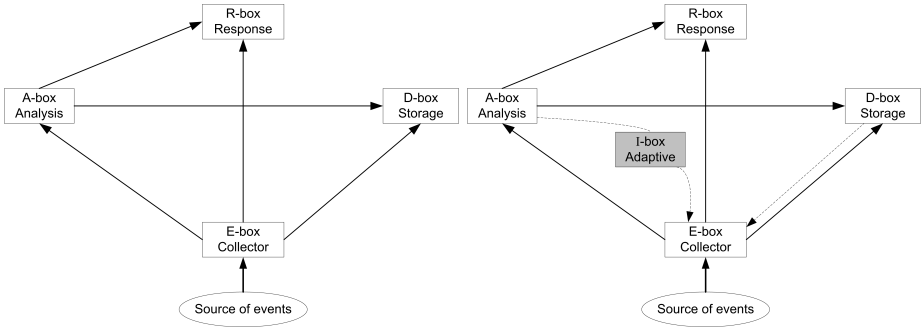

Figure 3: The new immunity-inspired framework architecture

tion will be based on danger signals. He proposed[1] an AIS based on DT ideas that is capable of handling the IDS alert correlation problems described above.

In our proposal (section 6) we believe that these signals, when analysed and correlated, can give important information from an ongoing attack and can be a key to in the definition of an efficient adaptive signature for the invader and/or attack.

6 A novel framework for Network Intrusion Detection

In this section we present a new framework for intrusion detection systems [4], which is an extension of the CIDF, detailed in [22, 8]. The CIDF was the result of an effort to develop tools and application programming interfaces so that intrusion detection research projects could evolve from a common reference and modular architecture. The CIDF models an IDS as an aggregate of four components or boxes (E-box, A-box, D-box and R-box) that inter-operate by processing, storing and signalling events.

Our framework has two major innovations: the definition of normal traffic, based on the data collected by the sensors and the deployment of immunological concepts to better cope with change and provide the system with some kind of memory of past events. This can be achieved by the adaptation of network sensors in order to better identify new event profiles and through the definition of a new box, the I-Box as illustrated in Figure 3. This I-Box implements the learning mechanisms of past events that should be stored in the D-Box for latter usage. The events generated by the A-box can also be used to generate new attack profiles, based on adaptive methodologies. This new component (I-box) can use immune algorithms to generate new event profiles (basically new cells) for the E-box, allowing the system to "learn" and better respond to future malicious attacks. This approach allows the IDS to "evolve" in an adaptive way and be self-adjusted by previously learnt attacks.

The meaning of "normal" and its distinction from abnormal traffic is a prob-

lem of very difficult solution. The approaches made so far through AIS have demonstrated some problems [23], mainly the high number of false positives events and the computational complexity related to its use in dynamic and large computer networks. We believe that the use of danger theory will allow us to contribute positively to a better definition of normality through the correlation of signals, thus decreasing significantly the number of false positives in behaviour-based network IDS.

7 Conclusions

In this paper we have presented and explained how some theoretical immunological concepts are being used in the development of a new generation of biologically inspired intrusion detection systems. We have shown some of the limitations found in the use of classical immunological concepts in the development of previous IDS and exposed a new and emergent Danger Theory (DT) as a new idea filled with potential strengths for the development of new adaptive and self-adjusted IDS. We have also described and proposed a new commomn framework for biologically inspired intrusion detection systmes that build on the previous and well established CIDF. We have also explained how we can take full advantage of the application of recent immunological concepts to better construct an IDS that we expect will be able to cope with some of the well known problems in current IDS methods for misuse and anomalies detection. We make effective of use network alert correlation techniques to characterize network traffic and define the meaning of what constitutes "normal" activity. We also intend to use alert correlation for danger signals based on DT and try to quantify the benefits that can be derived for an IDS in such an approach. We believe that the application of these emergent theoretical immunological concepts supported on DT will bring some good benefits to the deployment of IDS, by enabling the system with the capacity to better learning new and unforseen attacks in an adaptively and self-adjusted way.

References

[1] U. Aickelin, P. Bentley, S. Cayzer, J. Kim, and J. McLeod. Danger theory: The link between ais and ids? *Proc. of the Second Internation Conference on Artificial Immune Systems (ICARIS-03)*, pages 147–155, 2003.

[2] U. Aickelin and S. Cayzer. The danger theory and its application to artificial immune systems. *proceedings of The First International Conference on Artificial Immune Systems (ICARIS 2002)*, pages 141–148, 2002.

[3] U. Aickelin, J. Greensmith, and J. Twycross. Immune system approaches to intrusion detection-a review. *Proc. of the Second Internation Conference on Artificial Immune Systems (ICARIS-03)*, pages 316–329, 2004.

[4] M. Antunes, R. Monteiro, Correia M., and Santos H. Towards a new immunity-inspired intrusion detection framework. Submitted to The Second IEEE LCN Workshop on Network Security (WNS 2006), 2006.

[5] Y. Bai and H. Kobayashi. Intrusion detection systems: technology and development. *Advanced Information Networking and Applications, 2003. AINA 2003. 17th International Conference on*, pages 710–715, 2003.

[6] J. Balthrop, F. Esponda, S. Forrest, and M. Glickman. Coverage and generalization in an artificial immune system. *GECCO-2002: Proceedings of the Genetic and Evolutionary Computation Conference*, pages 3–10, 2002.

[7] G.R. Burmester and A. Pezzuto. *Color Atlas of Immunology*. Thieme Medical Publishers, 2003.

[8] Common Intrusion Detection Framework (CIDF). http://www.isi.edu/gost/cidf/, 2006.

[9] D. Dasgupta. An immune agent architecture for intrusion detection. *Proc. of GECCO 00–Workshop Proceedings*, pages 42–44, 2000.

[10] D. Dasgupta and F. Gonzalez. An immunity-based technique to characterize intrusions in computer networks. *Evolutionary Computation, IEEE Transactions on*, 6(3):281–291, 2002.

[11] D. Dasgupta and FA Gonzalez. An immunogenetic approach to intrusion detection. *Division of Computer Science, University of Memphis, Technical Report No. CS-01-001, May*, 2001.

[12] D. Dasgupta, N. Majumdar, and F. Nino. Artificial immune systems: A bibliography. *Computer Science Division, University of Memphis, Technical Report*, pages 02–001.

[13] L.N. de Castro. Immune engineering: Development of computational tools inspired in artificial immune systems. *(Portuguese), Ph. D. Thesis. DCA FEEC/UNICAMP, Campinas/SP. Brazil, May2001.*

[14] L.N. de Castro and J. Timmis. *Artificial Immune Systems: A New Computational Intelligence Approach*. Springer, 2002.

[15] FS de Paula, L.N. de Castro, and PL de Geus. An intrusion detection system using ideas from the immune system. *Evolutionary Computation, 2004. CEC2004. Congress on*, 1:1059–1066, 2004.

[16] H. Debar, M. Dacier, and A. Wespi. Towards a taxonomy of intrusion-detection systems. *Computer Networks*, 31(8):805–822, 1999.

[17] H. Debar and A. Wespi. Aggregation and correlation of intrusion-detection alerts. *Recent Advances in Intrusion Detection*, 2212:85–103, 2001.

[18] S. Forrest and S. Hofmeyr. Engineering an immune system. *Graft*, 4(5):5–9, 2001.

[19] S. Forrest, A.S. Perelson, L. Allen, and R. Cherukuri. Self-nonself discrimination in a computer. *Proceedings of the 1994 IEEE Symposium on Research in Security and Privacy*, pages 201–212, 1994.

[20] J. Greensmith, U. Aickelin, and J. Twycross. Detecting danger: Applying a novel immunological concept to intrusion detection systems. *6 th International Conference in Adaptive Computing in Design and Manufacture*, 2004.

[21] S.A. Hofmeyr and S. Forrest. Immunity by design: An artificial immune system. *Proceedings of the Genetic and Evolutionary Computation Conference (GECCO)*, pages 1289–1296, 1999.

[22] C. Kahn, P.A. Porras, S. Staniford-Chen, and B. Tung. A common intrusion detection framework. *Submitted to Journal of Computer Security, July*, 2000.

[23] J. Kim. *Integrating Artificial Immune Algorithms for Intrusion Detection.* PhD thesis, University of London, 2002.

[24] J. Kim and P. Bentley. An evaluation of negative selection in an artificial immune system for network intrusion detection. *Genetic and Evolutionary Computation Conference 2001*, pages 1330–1337.

[25] P. Matzinger. The danger model: A renewed sense of self. *Science's STKE*, 296(5566):301–305.

[26] P. Matzinger. http://cmmg.biosci.wayne.edu/asg/polly.html, 2006. The real function of the immune system by Polly Matzinger.

[27] C. Taylor and J. Alves-Foss. Nate: Network analysis of anomalous traffic events, a low-cost approach. *Proceedings of New Security Paradigms Workshop*, pages 89–96, 2001.

[28] R.E. Vance. Cutting edge commentary: A copernican revolution? doubts about the danger theory. *The Journal of Immunology*, (165):1725–1728, 2000.

[29] HS Venter and JHP Eloff. A taxonomy for information security technologies. *Computers and Security*, 22(4):299–307, 2003.