Vergleich von Lehransätzen für die Ausbildung in IT-Sicherheit

Inauguraldissertation
zur Erlangung des akademischen Grades
eines Doktors der Naturwissenschaften
der Universität Mannheim

vorgelegt von

Martin Mink
aus Offenbach am Main

Mannheim, 2009

Dekan:	Professor Dr. Ing. Felix Freiling, Universität Mannheim
Referent:	Professor Dr. Ing. Felix Freiling, Universität Mannheim
Korreferent:	Professor Dr. habil. Simone Fischer-Hübner, Karlstad University, Schweden

Tag der mündlichen Prüfung: 11. Dezember 2009

Kurzbeschreibung

Seit einigen Jahren werden in der universitären Ausbildung in IT-Sicherheit nicht nur Schutzmaßnahmen, sondern auch Angriffsmethoden vermittelt. Dies äußert sich in speziellen Praktika und auch in spielerischen Formen wie den so genannten Capture-The-Flag-Wettbewerben, in denen studentische Teams die Angreifer- und die Verteidigerrolle einnehmen. Der Ansatz, Angreifermethoden kennenzulernen, soll zu einem besseren Verständnis von IT-Sicherheit und deren Problemen führen, als das Vermitteln von reinen Schutzmaßnahmen. Die vom Autor mit diesem Ansatz seit dem Jahr 2003 an der RWTH Aachen und der Universität Mannheim gemachten Erfahrungen waren sehr positiv. So stellte sich die Frage, ob der Beitrag des Ansatzes zur Ausbildung von Sicherheitsexperten objektiv bewertet werden kann. Aber dazu muss Sicherheitswissen operationalisiert werden. Zu diesem Zweck wurde eine empirische Studie entworfen und durchgeführt, die die Auswirkung des Ansatzes auf das IT-Sicherheitsverständnis von Universitätsstudenten messen soll. Im Rahmen der Studie nehmen Studenten an einem entsprechend entworfenen Kurs zu IT-Sicherheit teil und mit geeigneten Messinstrumenten wird die Auswirkung des Ansatzes bestimmt. Bestandteil der Studie ist die Planung der Parameter der Studie, der Entwurf und die Durchführung von Kursen zur Informationssicherheit sowie der Entwurf von Tests, mit denen Sicherheitswissen gemessen werden kann.

In dieser Arbeit wird zum einen die Situation der IT-Sicherheitsausbildung an deutschen und internationalen Universitäten analysiert. Zum anderen wird die Planung und die Durchführung der Studie zur Bestimmung des Sicherheitsverständnisses vorgestellt. In der Studie wird der zu bewertende Lehransatz mit dem Lehransatz verglichen, der Schutzmaßnahmen in den Vordergrund stellt. Dafür wurde für jeden der Ansätze ein Einführungskurs in IT-Sicherheit mit theoretischen und praktischen Inhalten entwickelt. Bei der Auswertung der in den Studiendurchführungen empirisch erhobenen Daten zeigt sich eine Tendenz hin zu dem zu bewertenden Lehransatz, die allerdings nicht genug Aussagekraft hat. So müssen in weiteren Studiendurchführungen mehr Daten gesammelt werden, um eine Aussage treffen zu können.

Abstract

English title: Comparison of approaches for the education in information security

There's a tendency in IT security education at universities to not only teach defense mechanisms but also attack techniques. Increasingly more universities offer special labs, where students can experience both the attackers' and the security administrators' view. A more competitive form of such labs are the so-called Capture-The-Flag contests. Getting to know the attackers' side is thougt to lead to a better understanding of IT security and its problems compared to teaching only protective techniques. The authors' experiences with teaching attack techniques during his time at RWTH Aachen University and Mannheim University starting in 2003 were very positive. These experiences led to the question if it is possible to assess the contribution of the approach regarding the education of security professionals. This requires the IT security knowledge of students to be measured. For this purpose an empirical study was designed and conducted, with the goal of assessing the impact on the understanding of information security of university students. As part of the study, students attend a specially crafted course on IT security where adequate measures determine the effect of that approach. To conduct the empirical study, the study needs to be specified, the IT security courses will be created and conducted and tests to measure IT security knowledge will be designed.

This thesis analyzes the situation of IT security education at German and international universities. And it presents the design and the implementation of the mentioned empirical study. The study compares two approaches: the approach to be assessed and the one that focuses on defense mechanisms. To this end for each of the two approaches an introductory course on information security with theoretical and practical elements was designed. The evaluation of the empirical data gathered in the study showed a tendency in favor of the approach with focus on attack techniques, but the difference is not significant enough. To come to a conclusion regarding the assessment of the approach more empirical data has to be collected by repeating the study.

Danksagung

Bei der Durchführung meiner Arbeit habe ich viel Unterstützung erfahren. Bei all denjenigen möchte ich mir hier bedanken.

Insbesondere bei meinem Betreuer, Prof. Felix Freiling, für die Idee zum Thema der Arbeit und die Unterstützung bei der Durchführung; für Gespräche, Anregungen und Möglichkeiten. Dank auch an meine Kollegen des Lehrstuhls für Praktische Informatik I der Universität Mannheim für fruchtbare Diskussionen und Impulse.

Bei der Planung der empirischen Studie haben mich unterstützt

- ein Mitarbeiter des Instituts für Soziologie der RWTH Aachen, dessen Namen ich leider nicht mehr nachvollziehen kann, mit der Idee für die Konstruktion des Messinstruments in der Studie.
- Christian Spannagel von der PH Ludwigsburg mit Informationen zu empirischer Forschung und zur Durchführung einer empirischen Studie.
- Dr. Rainer Greifeneder vom Lehrstuhl für Mikrosoziologie und Sozialpsychologie der Universität Mannheim mit der Begleitung der zweiten Studiendurchführung, insbesondere Hilfe bei der Überarbeitung der Fragebögen und der Auswertung der Studie, sowie Durchsicht und Korrektur dieser Doktorarbeit.

Dank für die Hilfe bei der Durchführung der Studie gehen an

- die Mitarbeiter des Rechenzentrums der RWTH Aachen Thomas Paetzold (für Raum- und Infrastrukturnutzung), Frank Lindner und Sylvia Leyer (für die Netzwerkkonfiguration) sowie Jens Hektor.
- die Mitarbeiter des Rechnerbetriebs Informatik der RWTH Aachen Willi Geffers und Stefanie Scholten (für Raum- und Infrastrukturnutzung sowie Installation)
- Christian Ritter von der Universität Mannheim für die Überlassung des Rechnerpools der Informatik sowie für Ideen bezüglich und Umsetzung der Protokollierung.

- die studentischen Mitarbeiter Diego Biurrun (RWTH Aachen) für Unterstützung bei Installation und Administration während des Kurses sowie Ben Stock, Christoph Klasik, Christian Zimmermann und Christian Moch (Universität Mannheim) für die Unterstützung bei der Kursdurchführung. Ben hat außerdem das Kurssystem erstellt und die Auswertung unterstützt, Christoph und Christian Z. haben die Protokolle des Abschlusstest ausgewertet und Christian M. den Fragebogen der zweiten Studiendurchführung überarbeitet.

- Simon Rieche (RWTH Aachen) für Unterkunft, Fahrdienst und sonstige Unterstützung.

Besonders erwähnen möchte ich die Studenten der RWTH Aachen Frank van der Beek und Christian Mertens, die mir durch ihre Diplomarbeiten bei der Studienplanung und bei der Kursdurchführung eine große Hilfe waren.

Danken möchte ich Frau Professor Dr. habil. Simone Fischer-Hübner von der Universität Karlstad (Schweden) für die Übernahme des Zweitgutachtens.

Last but not least danke ich meiner Familie. Insbesondere meiner Frau Marion für ihre Geduld, für ihr Verständnis, für die Unterstützung meiner Arbeit, für ihre Bereitschaft, mehrere Male umzuziehen, für Anregungen und fürs Korrekturlesen. Sowie meinen Eltern Heidi und Stephan, dass Sie mir das Studium, und damit die Promotion, ermöglicht und mich in all der Zeit unterstützt haben.

Inhaltsverzeichnis

Tabellenverzeichnis

Abbildungsverzeichnis

1 Einleitung

In den letzten Jahren hat die Bedeutung der Informationstechnik (IT) einen immer größeren Stellenwert in der Gesellschaft eingenommen. Es ist kaum noch vorstellbar, wie öffentliche Einrichtungen oder große Unternehmen ohne den Einsatz der entsprechenden IT funktionieren. Die damit einhergehende fortschreitende Vernetzung der Computersysteme brachte nicht nur Vorteile mit sich, sondern erhöhte auch die Komplexität der Systeme sowie deren Anfälligkeit gegenüber Angriffen. Beispiele für Angriffe, die in der Vergangenheit besonders große Schäden verursachten, sind der im Jahr 2007 entdeckte Diebstahl von 45,7 Millionen Kreditkartennummern bei der US-amerikanischen Kaufhaus-Kette TJX [Spiegel Online, 2007], der Wurm Mydoom im Jahr 2004 [Heise online, 2004] und der Denial-of-Service-Angriff im Juli 2009 auf südkoreanische und US-amerikanische Regierungsseiten, Einkaufsportale und Nachrichtendienste [Heise online, 2009].

Mit dem rasanten Wachstum des Internets haben sich auch die Bedrohungen verändert. War beispielsweise noch vor einigen Jahren die Disketten die einzige Möglichkeit, Viren zu verbreiten, so werden heute innerhalb von Sekunden Millionen von E-Mails mit infiziertem Anhang verschickt. Des Weiteren haben sich auch die Zielsetzungen der Angreifer im Laufe der Zeit gewandelt. Viele Angreifer testen Systeme nicht mehr auf Schwachstellen, um ihren Ruf und ihr Ansehen in der Gemeinschaft zu steigern, sondern verfolgen zunehmend finanzielle Absichten.

Trotzdem investieren Unternehmen nur einen Bruchteil ihres IT-Budgets in IT-Sicherheit. Wie die jährlich veröffentlichte Studie des CSI [2008] aus dem Jahr 2008 feststellt, geben 53% der befragten Unternehmen weniger als 5% ihres gesamten IT-Budgets für IT-Sicherheit aus. Das Geld wird eher in neue Hardware und Software anstelle von IT-Sicherheitsschulungen für Mitarbeiter investiert. Schulungen für die Mitarbeiter sind aber

dringend notwendig, da in den meisten Fällen menschliches Fehlverhalten oder Unkenntnis die Ursache für Sicherheitsrisiken und -verletzungen ist, z. B. durch die Verwendung leicht zu ratender Passworte oder die Preisgabe von Informationen an Unberechtigte. Die Folgen sind eine hohe Zahl und eine lange Dauer von Systemausfällen, die bei kleineren und mittleren Firmen existenzbedrohende Ausmaße erreichen können. Die Ursache für das fehlende Risikobewusstsein ist häufig auf der Ebene des Managements zu finden, weil es Sicherheitsrisiken als technische und nicht als betriebswirtschaftliche Probleme betrachtet [Bräuer, 2006].

Das wesentlich schwerer erfassbare Sicherheitsrisiko ist der Anwender von IT-Systemen selbst. Wie der damalige Präsident des BSI, Herr Dr. Helmbrecht, auf dem IT-Sicherheitskongress des BSI im Jahr 2007 berichtete, mangelt es an einer Sozialisierung des Internets: Außerhalb des Internets ist es für jeden selbstverständlich, die Haustür zu schließen, wenn man das Haus verlässt, das Auto abzuschließen oder sein Geld auf die Bank zu bringen. In der virtuellen Welt machen sich die Benutzer aber nur wenige Gedanken, welche Konsequenzen ihr Handeln hat. Der Grund für dieses „naive“ Verhalten liegt unter anderem an der Risikowahrnehmung der Personen. Der amerikanische Kommunikationswissenschaftlers Peter Sandman hat dies einmal so zusammengefasst: „The risks that kill you are not necessarily the risks that anger and frighten you“ [Sandman, 1987]. So ist beispielsweise die Verwendung von Antiviren-Software weit verbreitet, weil hier das Risiko – u.a. geschürt durch Meldungen der Presse – als sehr hoch eingeschätzt wird. Anderen IT-Sicherheitsbereichen wie der Verschlüsselung von Datenbeständen oder Zugriffskontrollen wird hingegen weniger Beachtung geschenkt. Dieses so genannte *It won't happen to me-Syndrom* [Campbell u. a., 2007; Scheidemann, 2007] drückt aus, dass mit dem Glauben an die eigene Unverwundbarkeit bestimmte Risiken zu gering eingeschätzt werden, vor allem wenn bezüglich eines Risikos scheinbar noch kein Schaden eingetreten ist. In der Psychologie spricht man daher auch von einem „unrealistischen Optimismus“ [Scheidemann, 2007].

Letztendlich sind aber die meisten Ausfälle von Computersystemen nicht auf einen Ausfall von technischen Schutzmaßnahmen wie Firewalls oder Systeme zur Schutzzielverletzungserkennung (engl. Intrusion Detection Systems, IDS) zurückzuführen, sondern

eben auf menschliches Versagen, sei es durch Fehlverhalten von Mitarbeitern in einer bestimmten Situation, Fahrlässigkeit durch mangelndes Sicherheitsbewusstsein oder die nicht konsequente Einhaltung von Sicherheitsregeln und Vorschriften.

Vor allem in der Industrie ist der Wunsch nach mehr IT-Sicherheit sehr groß, allerdings ist die Anzahl qualifizierter Mitarbeiter gering. Mit ein Grund dafür ist, dass in der Hochschulausbildung zu wenig Informationssicherheit gelehrt und praktiziert wird und viele Studenten zu wenig Kenntnisse auf diesem Gebiet haben, wie auch eine von der Universität Regensburg im Jahr 2006 durchgeführte Studie belegt [Dimler u. a., 2006]. In dieser Studie stuften die befragten Studenten IT-Sicherheit zwar als sehr wichtig ein, es zeigte sich aber auch, dass ihnen Wissen und praktische Erfahrungen fehlten, um bestimmte Sachverhalte entsprechend zu beurteilen und in die Realität umzusetzen.

Um Schäden durch Sicherheitsvorfälle in der Zukunft besser verhindern zu können, muss in die *Aus- und Weiterbildung* von IT-Verantwortlichen und Anwendern investiert werden. Die vorliegende Arbeit konzentriert sich auf den Bereich der Aus- und Weiterbildung an Universitäten.

1.1 Stand der IT-Sicherheitslehre

In den letzten Jahren wurden verschiedene Angebote bezüglich Sicherheit in der Informationstechnik entwickelt. Viele Universitäten mit dem Studiengang Informatik bieten Vorlesungen zur IT-Sicherheit an, da dieses Thema aufgrund immer häufiger auftretender Meldungen über Sicherheitslücken sehr aktuell ist und man dringend gut ausgebildete Fachkräfte benötigt, die sich mit der Sicherheit von IT-Systemen auskennen. An Universitäten in Deutschland gibt es meistens nur einzelne Grundvorlesungen zu diesem Gebiet. Eine aktuelle Übersicht über Kurse in Deutschland, die sich mit IT-Sicherheit beschäftigen, scheint nicht zu existieren. Die von Jürjens [2005] stammt aus dem Jahr 2005. An ausgewählten amerikanischen Universitäten existieren seit dem Jahr 1999 so genannte *National Centers of Academic Excellence in Information Assurance Education* (CAEIAE) [NSA, 2009]. Dabei handelt es sich um von der National Security Agency (NSA) entworfene und unterstützte Programme mit dem Ziel die Schwachstellen in der

amerikanischen Informationsinfrastruktur zu reduzieren. Erreicht werden soll dieses Ziel durch die Förderung von IT-Sicherheitsausbildung und der Hervorbringung einer großen Anzahl gut ausgebildeter IT-Sicherheitsexperten.

Veranstaltungen und Kurse zu IT-Sicherheit werden in Deutschland von Institutionen wie Universitäten, Fachhochschulen, Berufsakademien und Unternehmen angeboten. Im Bereich der Universitäten gibt es zwei Ansätze der Lehre von IT-Sicherheit: zum einen als Studienschwerpunkt und zum anderen als Studiengang. Als Studienschwerpunkt wird IT-Sicherheit beispielsweise an der TU Darmstadt gelehrt, in dem man als Abschluss ein „Zertifikat IT-Sicherheit“ erwerben kann [Baier u. a., 2003]. Der erste Diplomstudiengang zu IT-Sicherheit in Deutschland existiert seit dem Jahr 2000 an der Ruhr-Universität Bochum, der durch einen Bachelor- und einen Masterstudiengang abgelöst wurde [Ruhr-Universität Bochum]. Ziel dieser Studiengänge ist die Ausbildung von Spezialisten für IT-Sicherheit.

International – insbesondere in den USA – werden viele Veranstaltungen und Kurse zu IT-Sicherheit angeboten, teilweise existieren auch spezielle Studiengänge. So bietet die University of Idaho seit 2006 einen „Master of Science in Information Assurance“ an [Taylor u. a., 2006]. Auch private Anbieter wie *EC-Council* [EC-Council] und *Offensive Security* [Offensive Security] bieten Ausbildung in IT-Sicherheit an.

Der Stand der IT-Sicherheitslehre an deutschen und internationalen Universitäten wird ausführlicher in Kapitel 2 vorgestellt.

1.2 Motivation

IT-Infrastrukturen werden von ständig wechselnden Gefahren bedroht. Ein Bericht von der „security threat trend front“ des Internets [Slewe u. Hoogenboom, 2004] zeigt eine Tendenz hin zu höherem Kenntnisstand der Angreifer und höherer Professionalität der Angriffe. Hacker wählen ihre Ziele mit wesentlich mehr Sorgfalt aus als zuvor und verwenden das komplette Waffenarsenal, wie mit Keyloggern ausgestattete Trojanische Pferde, um in Unternehmens- oder Verwaltungsnetzwerke einzudringen. Auf der Gegenseite verstärkt sich ebenfalls der Trend zu ganzheitlicher ausgerichteten Ansätzen. Dies bedeu-

tet, dass technische Lösungen wie Firewalls und Systeme zur Schutzzielverletzungserkennung in eine Sicherheits- und Risikomanagementperspektive integriert werden müssen. Trotzdem nennt Neumann, Gründer und Moderator des bekannten *Risks Digest* [Neumann], Management nicht als eine der acht wichtigsten aktuellen Herausforderungen von IT-Sicherheit [Neumann, 2004]: Neben praktischen Erfahrungen in Systementwicklung und Schutz der Privatsphäre (engl. „privacy“) wird das Thema Sicherheitsausbildung ebenfalls aufgeführt.

Die Sicherheitsausbildung an Universitäten ist geprägt von defensiven Techniken wie Kryptografie, Firewalls, Zugriffskontrolle und Erkennung von Schutzzielverletzungen. Aber auch hier lässt sich eine Tendenz in Richtung offensiver Methoden erkennen [Schumacher u. a., 2000; Vigna, 2003a]. Intuitiv sind offensive Methoden Verfahrensweisen, die das Ziel haben, etwas „kaputt zu machen“. Beispiele hierfür sind Daten im Netzwerk unbefugt mitzulesen (Bruch von Vertraulichkeit) oder den Zugriff auf einen Dienst durch einen Denial-of-Service-Angriff zu verhindern (Bruch von Verfügbarkeit). In der wissenschaftlichen Literatur erfahren offensive Techniken ebenfalls allgemeine Anerkennung [Arnett u. Schmidt, 2005; Farmer u. Venema, 1993; Arce u. McGraw, 2004]. Die Association for Computing Machinery (ACM) widmete sogar eine gesamte Sonderausgabe ihres Magazins „Communications of the ACM“ dem Thema „Hacking and Innovation“ [Conti, 2006].

Was ist der Grund dafür? In einem 2005 erschienenen Artikel argumentiert Conti [2005], dass Akademiker, die sich mit IT-Sicherheit beschäftigen, viel von Hackern und derem Sicherheitsdenken lernen können, indem sie deren Versammlungen besuchen, wie DEF CON [Defcon] oder Black Hat [BlackHat]. Dieser Ansatz stimmt mit der Entwicklung im professionellen Bereich überein, offensive Methoden in Sicherheitstests einzusetzen, speziell in der Variante Penetrationstest, welche den Einsatz von Hacker-Werkzeugen wie Netzwerk-Sniffern, Passwort-Crackern und Disassemblierern sowie von aktivem Testen von Unternehmensnetzwerken in Echtzeit beinhaltet.

Während der Lehrtätigkeit an der RWTH Aachen und der Universität Mannheim hat der Autor gute Erfahrungen mit der Verwendung von offensiven Methoden in Lehrveranstaltungen gemacht. Das Lehrangebot bestand primär aus einer Grundlagenvorlesung

zu IT-Sicherheit, einer praktischen Veranstaltung – dem so genannten Hacker-Praktikum –, verschiedenen Spezialvorlesungen (u.a. Sicherheit von Webapplikationen, Digitale Forensik) sowie Seminaren zum Thema IT-Sicherheit mit offensiven Anteilen. An den Lehrveranstaltungen nahmen regelmäßig viele Studenten teil – insbesondere an dem Hacker-Praktikum – und zeigten viel Begeisterung an IT-Sicherheitswettbewerben wie den so genannten Capture-The-Flag-Wettbewerben (siehe Abschnitt 1.4) – sowohl als Teilnehmer als auch als Zuschauer. Auch Unternehmen zeigten sich interessiert. Speziell nach Meldungen über Erfolge in Sicherheitswettbewerben oder über Forschungsergebnisse kamen Anfragen von Firmen nach Studenten mit Hacker-Erfahrung, nach Sicherheitsschulungen für die eigenen Mitarbeiter oder nach Wissen, um Firmen-Know-how zu schützen.

Betrachtet man die geschilderten Anzeichen, dann scheint es einen wesentlichen Vorteil zu geben, Sicherheit in offensiver Weise zu behandeln. Aber gibt es wirklich einen Vorteil? Und falls ja, ist dieser irgendwie quantifizierbar?

1.3 Beitrag

Basierend auf den Erfahrungen mit der Sicherheitsausbildung von Studenten entstand die Überzeugung, dass das Lehren von offensiven Methoden im universitären Lehrplan einen beachtlichen Vorteil hat. Kurz gefasst: In Lehrveranstaltungen sollte mehr Zeit auf den Angriff als auf die Verteidigung verwendet werden. Diese Feststellung führt jedoch zu einer grundlegenden Forschungsfrage: Lässt sich diese Hypothese objektiv messen? In dieser Arbeit wird ein experimenteller Aufbau beschrieben, mit dem die Hypothese „Angriff ist besser als Verteidigung“ überprüft werden soll. „Besser“ bedeutet in diesem Zusammenhang unter anderem, dass die Studenten durch die erlernten Angriffstechniken mehr Kenntnisse von potentiellen Schwachstellen in sicherheitskritischen Systemen haben und weniger Zeit benötigen, um sicherheitsbezogene Aufgaben ordnungsgemäß durchzuführen.

Die Hypothese wird durch eine empirischen Studie überprüft. Besser, als den offensiven Ansatz separat zu untersuchen, ist eine vergleichende Studie. Als Vergleich wurde die

klassische, defensiv geprägte Lehre gewählt. Für die Durchführung der Studie wurden zwei Kurse entworfen – der eine offensiv orientiert, der andere defensiv –, die ausgewählte Themen der IT-Sicherheit abdecken. Die Studie wurde geplant, durchgeführt und ausgewertet. Es wurde eine Metrik entworfen, um eine Bewertung des offensiven Ansatzes zu möglichen. An zwei Durchführungen der Studie nahmen insgesamt knapp 120 Studenten teil. Ergebnis der ersten Studiendurchführung war eine Tendenz zugunsten der offensiven Gruppe, in der zweiten Durchführung zugunsten der defensiven Gruppe. Die aufgestellt Vermutung lässt sich somit nicht belegen, ist aber auch nicht widerlegt. In weiteren Studiendurchführungen müssen mehr Daten erhoben werden, um eine Aussage treffen zu können.

1.4 Verwandte Arbeiten

Verschiedene Lehransätze beschäftigen sich mit offensiven Methoden in der Lehre. So wurde an der TU Darmstadt am Fachbereich Informatik seit 1999 regelmäßig ein „Hacker Contest“ genanntes Praktikum angeboten [Schumacher u. a., 2000]. Vergleichbare Praktika sind mittlerweile auch an anderen deutschen Universitäten Teil des Lehrangebots.

Es existieren weitere Projekte, die offensive Techniken als Lehrmethode einsetzen. Zu erwähnen sind hier die so genannten *Wargames*, die eine lange Tradition unter Sicherheitsinteressierten haben. Für ein Wargame erstellt ein Organisator eine Reihe von herausfordernden Aufgaben, die von den Teilnehmern gelöst werden müssen. Diese sind meist level-basiert und können – je nach Art – im Webbrowser oder auf der Kommandozeile bearbeitet werden. Die Aufgaben orientieren sich häufig an Problemen, die ein Angreifer bei dem Versuch einer Systemkompromittierung typischerweise überwinden muss (siehe auch Abschnitt 2.5.2). Etwas wettbewerbsorientierter als Wargames sind Capture-The-Flag- (CTF) oder Deathmatch-Wettbewerbe, in denen die teilnehmenden Teams versuchen, in die Computer der anderen Teams einzudringen, um so genannte *flags* (engl. für „Flaggen“) zu erobern und gleichzeitig den eigenen Server gegen Angriffe zu verteidigen. Am bekanntesten ist der „International CTF“ (iCTF) der University of

California at Santa Barbara (UCSB), in dem weltweit verteilte Teams gegeneinander antreten [iCTF]. Aber auch in Europa werden derartige Wettbewerbe organisiert, so z.B. CIPHER mit Ursprung an der RWTH Aachen [CIPHER]. Für mehr Informationen zu CTF-Wettbewerben siehe Abschnitt 2.5.2.

Im Militärbereich lassen sich ähnliche Ansätze offensiver Ausbildung finden, z.B. im US-amerikanischen Militär [White u. Nordstrom, 1998]. Das Information Technology and Operations Center der Militärakademie West Point (USA) verwendet ebenfalls offensive Methoden in seinem Lehrplan. Das Zentrum organisiert jährlich eine so genannte *Cyber Defense Exercise*, die Ähnlichkeiten mit den CTF-Wettbewerben hat. Einheiten der US-Streitkräfte mit einem Bereich in Sicherheitsausbildung wie die Militärakademie West Point, die U.S. Airforce Academy oder die Naval Postgraduate School nehmen an dieser Übung teil. Von den Teilnehmern betreute Computer werden von Mitarbeitern der National Security Agency (NSA) während einer Dauer von mehreren Tagen angegriffen und müssen von den Teilnehmern verteidigt werden [Schepens u. James, 2003; Dodge u. a., 2003].

Im Jahr 1997 wurde von Jonsson u. Olovsson [1997] ein Experiment durchgeführt, um das Verhalten eines Angreifers bei dem Versuch, ein System zu kompromittieren, zu analysieren und zu bewerten. Dazu wurden mehrere, voneinander unabhängige Teams gebildet, die die Rolle eines Angreifers übernahmen. Angriffsziel war ein Computersystem der Universität. Gemessen wurde die Zeit, die jedes Team an der Aufgabe arbeitet. In der Auswertung nach verstrichener Zeit zwischen zwei Angriffen ergaben sich zwei Cluster: ein Cluster mit (wenigen) Gruppen, die viel Zeit zwischen zwei Angriffen benötigten und ein Cluster mit den meisten Gruppen und kurzen Zeiträumen. Es wird die Vermutung aufgestellt, dass der Einbruchs-Prozess unterschieden werden kann in 1. eine Lernphase, 2. eine Standardangriffsphase und 3. eine innovative Angriffsphase.

Vigna [2003a] untersuchte die Lerneffekte von Veranstaltungen, in denen die Teilnehmer realistische Erfahrungen mit dem Prozess von Angriff und Verteidigung sammeln können. In Schritt 1 wurden die Teilnehmer in zwei Teams aufgeteilt, ein so genanntes *Red Team*, das für Angriffe und die Kompromittierung eines System zuständig war und ein *Blue Team*, dessen Aufgabe der Schutz des Systems und das Erkennen von Angrif-

fen war. In Schritt 2 mussten die beiden Teams sowohl angreifen als auch verteidigen (hierbei handelt es sich um den Vorläufer der oben genannten CTF-Wettbewerbe). In Schritt 3 – dem so genannten *Treasure Hunt* – mussten beide Teams in gegenseitiger Konkurrenz einen gezielten Angriff durchführen.

Näf u. Basin [2008] betrachten zwei Ansätze für praktische Kurse in Informationssicherheit: der Konflikt-basierte und der Begutachtungs-basierte. Sie stellen die Vermutung auf, dass jeder der beiden Ansätze aus vier Phasen besteht und sich die beiden Ansätze nur in einer Phase – der dritten – unterscheiden. Im Konflikt-basierten Ansatz schützen die Teilnehmer in dieser Phase ihr System und greifen gleichzeitig die Systeme der anderen Teilnehmer an, während im Begutachtungs-basierten Ansatz die Teilnehmer die von den anderen Teilnehmer implementierten Systeme begutachten. Näf u. Basin identifizieren durch die Untersuchung zwei prinzipielle Unterschiede der Ansätze: Zeitdruck und Perspektive. Im Konflikt-basierten Ansatz stehen die Teilnehmer unter wesentlich höherem Zeitdruck und befinden sich in der Rolle sowohl des Angreifers als auch des Verteidigers, während im Begutachtungs-basierten Ansatz nur geringer Zeitdruck besteht und nur die Rolle des Gutachters eingenommen wird. Bei der Untersuchung handelt es sich um einen konzeptuellen Vergleich des offensiven und des defensiven Ansatzes.

Keine der oben erwähnten Arbeiten versucht jedoch, den Vorteil der Verwendung von offensiven Methoden in der universitären Lehre abzuschätzen.

1.5 Gliederung

In Kapitel 2 wird eine Übersicht über die IT-Sicherheitsausbildung im Allgemeinen und an Universitäten gegeben. Kapitel 3 liefert den notwendigen Hintergrund für die durchgeführte empirische Untersuchung durch eine Einführung in empirische Methodik und in Metriken. Kapitel 4 stellt die Planung und die Durchführung der empirischen Studie vor. Die Ergebnisse der Studie werden in Kapitel 5 vorgestellt und bewertet. Kapitel 6 schließlich liefert eine Zusammenfassung und einen Ausblick.

1.6 Veröffentlichungen

Kapitel 1 enthält Teile von Mertens [2007], van der Beek [2007] und [Mink u. Freiling, 2006]. Kapitel 2 basiert teilweise auf der Diplomarbeit von Mertens [2007] sowie auf Dornseif u. a. [2005a], Dornseif u. a. [2005b] und Freiling u. Mink [2005]. Die Kapitel 3 – 5 basieren auf der Diplomarbeit von van der Beek [2007], die Kapitel 3 und 4 zusätzlich auf [Mink u. Freiling, 2006] und [Mink, 2007], Kapitel 3 auf [Mink u. Nowey, 2008] sowie die Kapitel 4 und 5 auf [van der Beek u. Mink, 2008]. Der Artikel [Mink, 2008] gibt einen Überblick über die durchgeführte Studie.

Nicht in diese Arbeit eingeflossen sind die folgenden Veröffentlichungen des Autoren: Ein Vergleich von Ansätzen zur Lehre von Digitaler Forensik [Dornseif u. a., 2006] (zusammen mit Philip Anderson, Maximillian Dornseif, Felix Freiling, Thorsten Holz, Alastair Irons und Christopher Laing), eine Fallstudie der Lehre von Digitaler Forensik [Freiling u. a., 2008] (zusammen mit Felix Freiling und Thorsten Holz) und eine Betrachtung von Datenschutz in Lehrveranstaltungen zu Digitaler Forensik [Liegl u. a., 2009] (zusammen mit Marion Liegl und Felix Freiling).

2 IT-Sicherheitslehre an Universitäten

Dieses Kapitel gibt einen Überblick über den derzeitigen Stand der IT-Sicherheitsausbildung an Universitäten. In Abschnitt 2.3 wird der Stand der IT-Sicherheitslehre an Universitäten anhand einer Untersuchung zu IT-Sicherheitsveranstaltungen vorgestellt. Zuvor stellt Abschnitt 2.2 grundlegende Strukturen in der IT-Sicherheitsausbildung und danach Abschnitt 2.4 Empfehlungen für die Ausbildung in IT-Sicherheit vor. Abschnitt 2.5 beschäftigt sich mit zwei grundlegenden Ansätzen in der Ausbildung – dem offensiv- und dem defensiv-orientierten Ansatz – und Abschnitt 2.6 geht auf Kritik am offensiven Ansatz ein.

2.1 Einleitung

Wie bereits deutlich gemacht, ist IT-Sicherheitsausbildung wichtig. Zur Abwehr von Schadsoftware wie Würmern und Botnetzen, von E-Mail-SPAM und von Angriffen auf Netzwerke und Webanwendungen, aber auch zur Awarenessbildung sind gut ausgebildete IT-Sicherheitsexperten nötig.

Die aktuelle Situation der IT-Sicherheitsausbildung in Deutschland zeigt eine im Jahr 2009 von der Hochschule Darmstadt im Auftrag von Bitkom und der Software AG durchgeführte Studie zur Informatikausbildung an deutschen Universitäten und Hochschulen [Knorz, 2008]. Im Rahmen der Studie wurden Studierende befragt. Ergebnis der Studie ist, dass der deutsche Informatik-Nachwuchs in Sachen IT-Sicherheit schlecht ausgebildet ist. Nur 37% der Studierenden haben vor dem Abschluss mehr als eine Veranstaltung mit IT-Sicherheitsinhalten besucht. Als Gründe werden die langsam mahlenden Mühlen der Hochschulen und die knappen Studienzeiten (insbesondere der Bachelor) angegeben. Die

Initiatoren der Studie fordern deswegen, Sicherheitsaspekte in die Lehre zu integrieren und verbindlich im Studienplan zu verankern. IT-Sicherheit solle als Querschnittaufgabe betrachtet werden: „Neun Minuten für die Sicherheit“ lautet die Forderung, d.h. in jeder Informatikvorlesung solle ein Anteil von 10% für Sicherheitsthemen verwendet werden (bezogen auf eine 90-minütige Veranstaltung). Es muss dann allerdings darauf geachtet werden, dass die Dozenten über das nötige Wissen und die entsprechende Motivation verfügen. Andernfalls könnten die Auswirkungen schlimmer sein, als überhaupt keine Sicherheit zu lehren.

Wie IT-Sicherheitsausbildung an Universitäten umgesetzt wird, unterscheidet sich von Land zu Land. In Deutschland werden die Konzepte der IT-Sicherheit in den Vorlesungen meist theoretisch erläutert. Zwar bieten die meisten Universitäten Übungen an, in denen die Studenten beispielsweise Fragen zu kryptografischen Verfahren beantworten müssen. Dies bedeutet aber, dass sie die in der Vorlesung vorgestellten Verfahren nicht praktisch beherrschen müssen, was dazu führt, dass die Studenten keine praktische Erfahrung für ihr späteres Berufsleben sammeln.

IT-Sicherheitsvorlesungen an internationalen Universitäten gehen im Hinblick auf die Vorbereitung für den späteren Berufsalltag häufig einen anderen Weg. Die Studenten müssen ebenfalls Übungsaufgaben bearbeiten, aber im Gegensatz zu den Aufgaben an deutschen Universitäten sind diese deutlich stärker an Problemen aus der Praxis orientiert. Des Weiteren werden vielfach so genannte „Labore“ (engl.: Labs) angeboten, in denen Studenten die Möglichkeit gegeben wird, praktische Aufgabenstellungen zu lösen. Ein weiterer Unterschied besteht darin, dass Studenten in den USA selbstständig Projekte aus dem Bereich IT-Sicherheit entwerfen und durchführen müssen. Das Resultat geht, wie die Bearbeitung der Hausaufgaben und die Note aus einer Klausur, in die Endnote mit ein. In den USA können bereits Undergraduate[1]-Studenten Vorlesungen zur IT-Sicherheit belegen. An deutschen Universitäten ist diese Möglichkeit in den meisten Fällen erst ab dem Hauptstudium gegeben.

Es werden *Einführungsveranstaltungen* in IT-Sicherheit von ausgewählten nationalen und internationalen Universitäten in Hinblick auf die gelehrten Themen untersucht und

[1]Dies entspricht dem Grundstudium eines deutschen Diplom-Studienganges.

basierend darauf eine Klassifizierung durchgeführt. Die Ergebnisse der Untersuchung und der Klassifizierung werden in Abschnitt 2.3 vorgestellt, zuerst jedoch im folgenden Abschnitt auf Strukturen in der IT-Sicherheitsausbildung eingegangen.

2.2 Strukturen in der IT-Sicherheitsausbildung

IT-Sicherheitsausbildung ist ein sehr komplexes und breit gefächertes Gebiet, welches nicht nur die IT-Infrastruktur, Software und sichere Programmierung umfasst, sondern unter anderem auch die aktuelle Gesetzeslage, ethische Grundlagen sowie eine permanente Weiterentwicklungen aufgrund rapider dynamischer Veränderungen. Es existieren drei große Bereiche, in denen eine fundierte IT-Sicherheitsausbildung sehr wichtig ist [Highland, 1992]:

Industrie: Unternehmen, die in die Sicherheit von Informationen investieren, stellen häufig nur die technischen Aspekte in den Vordergrund. Dabei wird der entscheidende Faktor oftmals nicht beachtet, denn nicht die Technik allein ist wichtig, um Informationssicherheit zu gewährleisten, sondern auch die Personen, die mit der Technik arbeiten. Da sich ein potentieller Angreifer immer das schwächste Glied aussucht, können keine IT-Systeme sicher betrieben werden, solange nicht alle Anwender ausreichend geschult und über mögliche Gefahrenquellen oder Risiken aufgeklärt wurden [Yngström u. Björck, 1999]. Dabei ist gerade in mittelständischen oder kleinen Firmen die IT-Sicherheitsausbildung der Mitarbeiter sehr wichtig, da diese im Falle eines Systemfehlers in der Regel nicht über geeignete Ausweich- oder Backupsysteme verfügen.

Akademischer Bereich: Auch an Hochschulen mangelt es an Fachkräften, um die komplexen und vielseitigen Inhalte von IT-Sicherheit zu lehren. Im Jahr 2006 empfahl die Gesellschaft für Informatik (GI), IT-Sicherheit in der schulischen und akademischen Ausbildung aller Studiengänge im Curriculum stärker zu berücksichtigen. Auch die Möglichkeit der Einführung eines spezialisierten Studiengangs „Informationssicherheit“ wird diskutiert [Gesellschaft für Informatik e.V., 2006].

Regierung: IT-Sicherheitsexperten sind auch innerhalb der Regierung gefragt, z.B. zum Schutz von nationalen Interessen wie für den Aufbau eines Frühwarnsystems für das Internet [CarmentiS]. Auch bieten sich Strafverfolgungsbehörden durch die zunehmende Vernetzung immer weniger reale Ermittlungsansätze, sodass mehr auf dem virtuellen Weg ermittelt werden muss. Dies betrifft auch die Entwicklung und den Einsatz der so genannten *Remote-Forensic-Software* der Bundesregierung, besser bekannt als *Bundestrojaner*, zur Durchführung von Onlinedurchsuchungen. Dabei sind vor allem Programmierer mit viel Kenntnissen von offensiven Techniken gefragt – auch um sicherzustellen, dass dieses Programm weder Sicherheitslücken verursacht noch durch Dritte missbraucht werden kann.

Das US-amerikanische *National Institute of Standards and Technology* (NIST) strukturiert die Lehre in Informationssicherheit in drei Ebenen, wie sie Abbildung 2.1 zeigt. Die unterste Ebene ist die *Awareness*-Schicht, in der die grundlegenden Verhaltensrichtlinien in der IT-Sicherheit vermittelt werden sollen und die sich daher an alle Personen richtet, die mit Informationssystemen arbeiten. Die zweite Schicht bildet das *Training*, welches nur die Personen erhalten sollen, die ein spezielles Wissen über mögliche Bedrohungen und Risiken für die weitere Ausübung ihrer aktuellen Tätigkeit aufweisen müssen. Die oberste Schicht wird durch den Begriff der *Education* geprägt, welche primär auf diejenigen Personen ausgerichtet ist, die IT-Sicherheit beruflich ausüben [de Zafra u. a., 1998]. Im Folgenden werden diese drei Schichten genauer beschrieben.

Abbildung 2.1: Die drei Lehrebenen nach NIST.

2.2.1 Awareness

Mit *Awareness* wird die Einstellung einer Person oder Gemeinschaft gegenüber einem bestimmten Sachverhalt bezeichnet [de Zafra u. a., 1998]. *Security Awareness* steht hier also für das Verhalten bzw. das entwickelte Bewusstsein gegenüber Themen der Informationssicherheit. IT-Sicherheit ist ein wichtiges Thema, das an eine breite Öffentlichkeit adressiert ist und für jeden, der privat oder geschäftlich mit einem Computer arbeitet, relevant ist. Daher gehört Awareness zu einem besonders kritischen Punkt, welcher die Grundlage und Basis jeder weiterführenden Ausbildung und tiefer gehenden Lehre bildet. In vielen großen Unternehmen existieren bereits spezielle Awarenesskampagnen, um die Sensibilität der Mitarbeiter gegenüber sicherheitskritischen Aspekten aufzubauen und den Sinn von Sicherheitsrichtlinien näher zu bringen.

Awarenesskampagnen müssen jedoch immer aktuell gehalten werden, um die Zielgruppe immer wieder von neuem anzusprechen bzw. zu motivieren. Das können spezielle Lehrvideos oder neuartige Konzepte wie ein Infotainmentsystem sein, wie es beispielsweise bei Airbus Industries realisiert wurde [Aust, 2007], um den Angestellten wesentliche Aspekte der Informationssicherheit auf spielerische Art und Weise zu vermitteln und dadurch den Lerneffekt zu erhöhen. Besonders das Auslösen persönlicher Betroffenheit wirkt dem in Kapitel 1 vorgestellten *It won't happen to me*-Syndrom entgegen, was die Erfolgswahrscheinlichkeit von Awarenesskampagnen zusätzlich erhöht [Scheidemann, 2007].

Awarenessaktivitäten sind vor allem dadurch charakterisiert, dass es sich um sehr kurze Maßnahmen handelt, durch welche die Adressaten der Kampagne neu aufgenommene Erfahrungen in ihr bestehendes Verhalten, also ihre täglichen Arbeitsabläufe, integrieren sollen. Dabei soll dieses Ziel ausschließlich durch die Vermittlung von Informationen aus dem Bereich IT-Sicherheit erreicht werden, nicht durch technische Details oder die zur Realisierung von Sicherheit benötigten Technologien.

Es muss jedoch berücksichtigt werden, dass es nicht unbedingt einen Zusammenhang zwischen Wissen und Verhalten gibt. Die so genannte „Kluft zwischen Wissen und Handeln“ führt zum Beispiel dazu, dass Menschen rauchen, obwohl sie sich der Gesundheitsschädlichkeit bewusst sind. Laut einer Meldung auf Heise online [2008] existieren

Studien und Umfragen, die belegen, dass es keine direkte Verbindung zwischen IT-Wissen und IT-Verhalten gibt, es ließ sich jedoch keine derartige Studie ermitteln.

2.2.2 Training

Eine Trainingsveranstaltung ist formaler aufgebaut als eine Awarenesskampagne und verfolgt das Ziel, bei den Teilnehmern Wissen und Fähigkeiten (engl.: Skills) aufzubauen. Awarenessprogramme fallen nicht in die Kategorie des Trainings, sondern bilden eine Vorstufe, da alle im Training angeeigneten Fähigkeiten auf Awareness aufbauen. So wäre im Bereich der Passwortsicherheit ein kleiner Aufkleber auf der Tastatur mit der Aufforderung, keine kurzen Passwörter zu verwenden, eine sehr einfache Awarenessaktivität, während im Training genauere Details zur Struktur von Passwörtern, Parametern und Passwortänderungen behandelt werden. Eine Trainingseinheit richtet sich somit nicht mehr an alle PC-Benutzer, sondern an speziell ausgewählte Personen, die bestimmte Fähigkeiten und Kenntnisse für ihre tägliche Arbeit benötigen und dadurch ihre Arbeitsleistung verbessern können. Ein Beispiel für eine Trainingsaktivität ist ein IT-Sicherheitskurs für Administratoren, in welchem bestimmte Themen wie Management, Kryptografie, Support und logische sowie physische Zugriffskontrollen vertieft behandelt werden.

Bestandteil einer Trainingsveranstaltung ist allerdings nicht, *warum* eine durchgeführte Handlung zur Erhöhung der IT-Sicherheit beiträgt oder wie spezielle Algorithmen einer Sicherheitssoftware funktionieren. Primäres Lehrziel ist hier, *wie* eine bestimmte Fragestellung beziehungsweise ein Problem zu lösen ist. Die Vermittlung tiefergehender Kenntnisse ist die Aufgabe der Education-Schicht.

2.2.3 Education

Die Education-Ebene bildet die oberste Schicht der Lehrpyramide und zugleich das Hauptproblem in der IT-Sicherheitsausbildung. Während bei einer Trainingsaktivität nur benötigte Kenntnisse für eine bestimmte Tätigkeit gelehrt werden, zielt Education darauf ab, die vermittelten Fähigkeiten aus allen Trainingseinheiten zu einem ge-

meinsamen Wissensteil zu kombinieren [Wilson u. Hash, 2003]. Dazu gehört es auch, fächerübergreifend auf verschiedene Konzepte und Probleme einzugehen und diese zu analysieren, um einerseits proaktiv in bestimmten Situationen reagieren zu können, bevor es zu Beeinträchtigungen wie Datenverlusten oder Systemausfällen kommt, oder andererseits um gelernte Kenntnisse schneller und leichter auf unbekannte Sachverhalte und Probleme übertragen zu können.

Ein Beispiel für akademische Education ist der Diplomstudiengang Informatik, in welchem Studenten eine Vorlesung oder mehrere besuchen, um neue Fähigkeiten zu erlernen oder vorhandene Kenntnisse zu vertiefen. Die akademische Education schließt dabei nicht nur die Lehre für Studenten, sondern auch die weiterführende Ausbildung an der Universität nach dem Studienabschluss mit ein. So unterscheidet man hinsichtlich der Inhalte und der Struktur zwischen Education für Studenten und für Absolventen.

Education für *Studenten* bezieht sich auf die Ausbildung im Grund- und Hauptstudium, in welcher kein Fokus auf bestimmte Themenbereiche gelegt wird. Stattdessen wird eine Fülle von Informationen aus dem gesamten IT-Bereich vermittelt, um den Studenten ein kompaktes Wissen für verschiedene Informatikbereiche zu vermitteln.

Education für *Absolventen* bezieht sich zum Beispiel auf die Promotion von Studenten nach ihrem Studium zur Erlangung des Doktorgrades. Im Gegensatz zur studentischen Education liegt hier der Fokus auf einem bestimmten Spezialgebiet, z.B. WLAN-Sicherheit, welches über mehrere Jahre wissenschaftlich bearbeitet und untersucht wird.

Tabelle 2.1 fasst die Unterschiede zwischen den drei Lehrebenen zusammen.

2.3 Klassifikation von IT-Sicherheitsveranstaltungen

Wir stellen nun die Ergebnisse einer Studie vor, die Einführungsveranstaltungen in Informationssicherheit an Universitäten im Hinblick auf die gelehrten Themen untersucht [Mertens, 2007]. Die Studie wurde 2007 im Rahmen einer vom Autor betreuten Diplomarbeit durchgeführt. Für die Untersuchung wurden 18 deutsche und 11 internationale Universitäten ausgewählt, die entsprechende Veranstaltungen anbieten. Für die Auswahl der deutschen Universitäten wurde das vom Centrum für Hochschulentwicklung (CHE)

	Awareness	**Training**	**Education**
Grundfrage	Was ist wichtig?	Wie erreicht man etwas?	Warum ist das so?
Was wird vermittelt?	Information	Wissen	Erkenntnis
Ziel	Erkennen und Beibehalten	Fertigkeiten aufbauen	Verständnis
Lehrmethode	Beispiele, Videos, Poster	Praxis und Vorlesungen, Fallstudien	Theoretische Diskussionen, Seminare, Forschungen
Überprüfung	Tests – Wiedergabe des Gelernten	Praxis – Anwendung des Gelernten	Ausarbeitung – Interpretation des Gelernten
Dauer	Kurz	Mittel	Lang

Tabelle 2.1: Gegenüberstellung der drei Lehrebenen

durchgeführte Hochschulranking des Jahres 2006 für das Fach Informatik zugrunde gelegt [Centrum für Hochschulentwicklung, 2006]. Die internationalen Universitäten wurden aus dem „Academic Ranking of World Universities – 2005“ [Institute of Higher Education, 2005]) ausgewählt. Die Liste wurde ergänzt durch Teilnehmer des Capture-The-Flag-Wettbewerbs iCTF [iCTF] aus dem Jahr 2006.

Mithilfe eines Fragebogens wurden von dem jeweiligen Dozenten Informationen zur Veranstaltung erhoben, falls sie nicht aus öffentlich zugänglichen Quellen ermittelbar waren. Dies beinhaltete speziell die Vorlesungsinhalte, aber auch weitere Informationen, beispielsweise wie oft die Veranstaltung bereits stattgefunden hat, verwendete Literatur und Altersklasse des Dozenten. Mithilfe einer Clusteranalyse basierend auf den gelehrten Themen wurden drei Cluster identifiziert (siehe Tabelle 2.2), die im folgenden vorgestellt werden. In zwei Fällen (nämlich Universität Mannheim/RWTH Aachen und TU Berlin/TU Ilmenau) wurden zwei Universitäten zusammen betrachtet, da der jeweilige Dozent die Universität gewechselt hatte.

Für die Clusteranalyse wurde als Proximitätsmaß der Jaccard-Koeffizient, d.h. ein Ähnlichkeitsmaß, verwendet. Anschließend wurde der Single-Linkage-Algorithmus als

Fusionierungsalgorithmus auf die resultierende Ähnlichkeitsmatrix angewandt, um Ausreißer zu identifizieren. Aufgrund dieser Auswertung wurden die TU Dresden, die TU Wien und die University of Illinois at Urbana-Champaign als Ausreißer angesehen und aus den folgenden Betrachtungen ausgeschlossen. Für mehr Informationen zu Clusteranalyse siehe [Backhaus u. a., 2006; Bacher, 1994; Steinhausen u. Langer, 1977; Hudec, 2003]

Es sei darauf hingewiesen, dass sich die Aussagen auf *Einführungsveranstaltungen* beziehen und es sich nicht um allgemein gültige Aussagen handelt. Eine dem konservativen Cluster zugeordnete Universität kann sehr wohl weitere Veranstaltungen anbieten, die offensiv orientiert sind (z.B. ein Praktikum).

2.3.1 Untersuchte Themen

Bevor genauer auf die Ergebnisse der Clusteranalyse eingegangen wird, werden in Tabelle 2.3 die Häufigkeiten der einzelnen Themen angegeben, um zu sehen, welche Themen am häufigsten gelehrt werden. In der zweiten Spalte der Tabelle steht die absolute Anzahl des jeweiligen Themas und in der dritten Spalte der prozentuale Anteil, der angibt, in wie viel Prozent der untersuchten Vorlesungen dieses Thema gelehrt wird.

Aus Tabelle 2.3 lässt sich ablesen, dass Themen aus dem Bereich Kryptografie, wie symmetrische und asymmetrische Verschlüsselung (89,7%), kryptografische Hashfunktionen (86,2%) und Kryptoanalyse (72,4%), in einem Großteil der untersuchten Vorlesungen vorkommen. Ebenfalls auf den oberen Rängen stehen die Themen Passwortsicherheit (75,9%) und Zertifikate (82,8%).

Weniger Berücksichtigung durch die Dozenten erfahren Themen wie Cross-Site Scripting, Portscanning und anonymes Surfen, die jeweils nur in 24,1% der untersuchten Kurse behandelt werden. Des Weiteren sind auch die Themen SQL-Injection, Race Conditions und Sicherheit in mobilen Netzwerken nur schwach, d. h. ungefähr nur in jeder vierten Vorlesung, vertreten. Ein möglicher Grund für die schwache Repräsentation dieser Themengebiete könnte sein, dass diese erst in den letzten Jahren mit der stärkeren Verbreitung des Internets in den Mittelpunkt der IT-Sicherheit traten. Die kryptografischen Verfahren hingegen existieren schon seit Jahrzehnten und sind deshalb in den Curricula

Tabelle 2.2: Ergebnis der Clusteranalyse

Name der Universität	**Zugeordnetes Cluster**
Univ. Mannheim/RWTH Aachen	1
Univ. Potsdam	1
Univ. Dortmund	1
Univ. Karlsruhe	1
Uppsala University	2
Univ. Saarbrücken	2
Univ. of South Florida	2
Univ. Regensburg	2
Imperial College London	2
FU Berlin	2
Columbia Univ.	2
North Carolina State	2
Pennsylvania State	2
TU München	2
UCSB	2
TU Darmstadt	2
Universität Zürich	2
Polytechnic Univ., Brooklyn	2
Univ. Rostock	2
Univ. Ulm	2
Univ. Magdeburg	2
Univ. of Colorado Springs	2
Univ. Lübeck	2
Univ. Freiburg	2
Univ. Erlangen	3
TU Berlin/TU Ilmenau	3
Yale Univ.	3
TU Kaiserslautern	3
Univ. Hamburg	3

Tabelle 2.3: Häufigkeiten der Themen

Thema	**absolute Häufigkeit**	**prozentuale Häufigkeit**
Symmetrische/Asymmetrische Verschlüsselung	26	89,7
Kryptografische Hashfunktionen	25	86,2
Passwort-Sicherheit	22	75,9
Zertifikate	21	82,8
Kryptoanalyse	21	72,4
Digitale Signaturen	20	69,0
Firewalls	19	65,5
Viren, Trojaner, Würmer	19	65,5
IPSec	17	58,6
Denial of Service-Angriffe	17	58,6
SSL/TLS	17	58,6
Sicherheits-Modelle	17	58,6
Email-Sicherheit	17	58,6
Buffer Overflows	16	55,2
Zugriffskontrollstrategien	16	55,2
IP-, ARP-, DNS-Spoofing	15	51,7
Intrusion Detection Systeme	13	44,8
OS-Sicherheit	13	44,8
Virtual Private Networks (VPN)	12	41,4
Sicherheitskriterien	11	37,9
Bedrohungs- und Risikoanalyse	10	34,5
Sniffing	10	34,5
Sandbox	9	31,0
SQL-Injection	8	27,6
Sicherheit in mobilen Netzwerken	8	27,6
Race Conditions	8	27,6
Cross-Site Scripting	7	24,1
Portscanning	7	24,1
Anonymes Surfen	7	24,1

verankert.

Die Resultate der Clusteranalyse zeigen, dass es insgesamt drei Cluster gibt, wovon zwei eine relativ geringe Anzahl von Objekten enthalten. Die Entscheidung, welche Merkmale, also Themen, die einzelnen Cluster am besten repräsentieren wurde mit Hilfe von Häufigkeiten realisiert. Dazu wurde die Häufigkeit jedes Themas in jedem Cluster ermittelt und dann analysiert, welche Themen am häufigsten auftreten.

2.3.2 Das innovative Cluster

Für das erste Cluster, bestehend aus der Universität Mannheim/RWTH Aachen, der Universität Potsdam, der Universität Karlsruhe und der Universität Dortmund ergibt sich die in Abbildung 2.2 dargestellte Tabelle, welche Aufschluss über die gelehrten Themen dieses Clusters gibt.

Man erkennt, dass folgende Themen für diese Gruppe typisch sind, da sie in jeder der Vorlesungen gelehrt werden: Race Conditions, Buffer Overflow, Sniffing, Denial of Service Angriffe, Portscanning, Cross-Site Scripting, Viren, Würmer, Trojaner und Passwort-Sicherheit. Diese Themen stehen noch nicht lange im Mittelpunkt der IT-Sicherheitslehre, da sie z.T. erst in den letzten Jahren an Aktualität gewonnen haben. Dagegen werden Themen wie Kryptoanalyse, Sandbox, Intrusion Detection Systeme, Anonymes Surfen, Digitale Signaturen und Sicherheit in mobilen Netzwerken, die einen eher „defensiven“ Charakter aufweisen, weniger oder gar nicht berücksichtigt. Da dieses Cluster zum größten Teil *aktuelle* Themen der IT-Sicherheit behandelt, wird es als die „innovative“ Gruppe bezeichnet.

2.3.3 Das ausgewogene Cluster

Das zweite Cluster – für die enthaltenen Universitäten siehe Tabelle 2.2 – ist eine gemischte Gruppe, da es bis auf das Thema symmetrische und asymmetrische Verschlüsselung, welches in jeder Vorlesung behandelt wird, alle Gebiete zu einem gewissen Prozentsatz enthält, siehe Abbildung 2.3. Eine eindeutige Tendenz zu einem bestimmten Themengebiet ist nicht erkennbar, da einerseits die Themen aus dem Gebiet der

Statistics

	Anzahl	Prozentzahl
Kryptoanalyse	4	,00
Sniffing	4	100,00
Race Conditions	4	100,00
IP, ARP, DNS Spoofing	4	75,00
Denial of Service Angriffe	4	100,00
Sandbox	4	,00
Buffer Overflow	4	100,00
SQL-Injection	4	50,00
Cross-Site-Scripting	4	100,00
Viren, Trojaner, Würmer	4	100,00
Portscanning	4	100,00
Intrusion Detection Systeme	4	,00
Virtual Private Network	4	25,00
Sicherheitskriterien	4	25,00
Bedrohungs- und Risikoanalyse	4	25,00
Kryptografische Hashfunktionen	4	25,00
Sicherheits-Modelle	4	50,00
Passwort-Sicherheit	4	100,00
Zugriffskontrollstrategien	4	25,00
OS-Sicherheit	4	75,00
Symmetrische/ Asymmetrische Verschlüsselung	4	25,00
IPSec	4	25,00
Anonymes Surfen	4	,00
Firewall	4	75,00
Digitale Signaturen	4	,00
Zertifikate	4	25,00
SSL/TLS	4	25,00
Sicherheit in Mobile Networks	4	,00
Email-Sicherheit	4	25,00

Abbildung 2.2: Häufigkeitstabelle für die Themen des innovativen Clusters

Kryptografie, wie Kryptoanalyse (80%), kryptografische Hashfunktionen (95%), Digitale Signaturen (90%) und Zertifikate (90%) und andererseits die Themen Email-Sicherheit (80%), Passwort-Sicherheit (80%) und Viren, Würmer und Trojanern (75%) vertreten sind. Wenig berücksichtigt werden in diesen Veranstaltungen die Themen Portscanning (10%), Cross-Site Scripting (15%) und Race Conditions (20%). Da in diesem Cluster ein ausgewogenes Verhältnis zwischen den untersuchten Themengebieten besteht, wird es als die „ausgewogene“ Gruppe bezeichnet.

2.3.4 Das konservative Cluster

Die dritte und letzte Gruppe, dargestellt in Abbildung 2.4, besteht aus den fünf Universitäten Erlangen, Berlin/Ilmenau, Yale, Kaiserslautern und Hamburg. Es ist ersichtlich, dass in den Vorlesungen dieser Universitäten der Fokus auf Themen aus dem Bereich Kryptografie liegt, da die Bereiche Kryptoanalyse, kryptografische Hashfunktionen, symmetrische und asymmetrische Verschlüsselung, Zertifikate und SSL/TLS jeweils in allen Vorlesungen dieser Kategorie behandelt werden. Nicht angesprochen werden Themen, die mit der Sicherheit von Netzwerken oder Software zu tun haben, wie z. B. Sniffing und Buffer Overflows. Aufgrund des deutlichen Gegensatzes zur ersten, der innovativen Gruppe, wird dieses Cluster als die „konservative“ Gruppe bezeichnet.

Zusammenfassend kann man die drei Partitionen folgendermaßen charakterisieren: Das erste Cluster, die innovative Gruppe, verzichtet größtenteils auf das Lehren von kryptografischen Grundlagen und behandelt stattdessen aktuelle Themen der IT-Sicherheit. Das dritte Cluster kann man als die konservative Gruppe bezeichnen, da hier hauptsächlich Kryptografie gelehrt wird. Die ausgewogene Gruppe steht, wie der Name schon sagt, für die Vorlesungen, in denen eine ausgewogene Mischung der Themen behandelt wird.

Interessant ist die Frage, in welche Cluster Universitäten mit Teams, die an Capture-The-Flag-Wettbewerben (siehe Abschnitt 2.5.2) teilnehmen, bei der Clusteranalyse eingeordnet werden. Bis auf drei Ausnahmen, nämlich Universität Mannheim/RWTH Aachen, TU Berlin und Universität Hamburg, wurden alle Teilnehmer dem ausgewogenen Cluster zugeordnet. Die TU Berlin sowie die Universität Hamburg wurden dem

Statistics

	Anzahl	Prozentzahl
Kryptoanalyse	20	80,00
Sniffing	20	30,00
Race Conditions	20	20,00
IP, ARP, DNS Spoofing	20	55,00
Denial of Service Angriffe	20	60,00
Sandbox	20	45,00
Buffer Overflow	20	60,00
SQL-Injection	20	30,00
Cross-Site-Scripting	20	15,00
Viren, Trojaner, Würmer	20	75,00
Portscanning	20	10,00
Intrusion Detection Systeme	20	60,00
Virtual Private Network	20	40,00
Sicherheitskriterien	20	50,00
Bedrohungs- und Risikoanalyse	20	45,00
Kryptografische Hashfunktionen	20	95,00
Sicherheits-Modelle	20	75,00
Passwort-Sicherheit	20	80,00
Zugriffskontrollstrategien	20	65,00
OS-Sicherheit	20	50,00
Symmetrische/ Asymmetrische Verschlüsselung	20	100,00
IPSec	20	60,00
Anonymes Surfen	20	35,00
Firewall	20	70,00
Digitale Signaturen	20	90,00
Zertifikate	20	90,00
SSL/TLS	20	55,00
Sicherheit in Mobile Networks	20	25,00
Email-Sicherheit	20	80,00

Abbildung 2.3: Häufigkeitstabelle für die Themen des ausgewogenen Clusters

Statistics

	Anzahl	Prozentzahl
Kryptoanalyse	5	100,00
Sniffing	5	,00
Race Conditions	5	,00
IP, ARP, DNS Spoofing	5	20,00
Denial of Service Angriffe	5	20,00
Sandbox	5	,00
Buffer Overflow	5	,00
SQL-Injection	5	,00
Cross-Site-Scripting	5	,00
Viren, Trojaner, Würmer	5	,00
Portscanning	5	20,00
Intrusion Detection Systeme	5	20,00
Virtual Private Network	5	60,00
Sicherheitskriterien	5	,00
Bedrohungs- und Risikoanalyse	5	,00
Kryptografische Hashfunktionen	5	100,00
Sicherheits-Modelle	5	,00
Passwort-Sicherheit	5	40,00
Zugriffskontrollstrategien	5	40,00
OS-Sicherheit	5	,00
Symmetrische/ Asymmetrische Verschlüsselung	5	100,00
IPSec	5	80,00
Anonymes Surfen	5	,00
Firewall	5	40,00
Digitale Signaturen	5	40,00
Zertifikate	5	100,00
SSL/TLS	5	100,00
Sicherheit in Mobile Networks	5	60,00
Email-Sicherheit	5	,00

Abbildung 2.4: Häufigkeitstabelle für die Themen des konservativen Clusters

konservativen Cluster und die Universität Mannheim/RWTH Aachen dem innovativen Cluster zugeordnet. Interessant ist das Ergebnis, dass sich alle *internationalen* Teilnehmer des CTF-Wettbewerbs in der ausgewogenen Klasse befinden. Dies deutet darauf hin, dass in den IT-Sicherheitsvorlesungen dieser Universitäten auf die Lehre von aktuellen Themen kombiniert mit Kryptografie geachtet wird.

In den folgenden drei Abschnitten werden nun noch die drei ermittelten Cluster im Hinblick auf das Alter des Dozenten, die verwendete Literatur und die Häufigkeit, wie oft eine Vorlesung stattfand, untersucht.

2.3.5 Alter der Dozenten

Eine interessante Frage ist, ob man bestimmten Altersgruppen spezielle Themen zuordnen kann, z. B. ob ältere Dozenten eher Kryptografie oder die Funktionsweise von Firewalls als aktuelle Themen lehren. Um dieser Frage nachzugehen ist in Tabelle 2.4 die Altersstruktur aller Dozenten der untersuchten Vorlesungen dargestellt, in der die – in der Clusteranalyse als Ausreißer betrachteten – Technische Universität Wien, TU Dresden und University of Illinois at Urbana-Champaign enthalten sind. Es ist anzumerken, dass an der Universität Hamburg und an der TU Wien jeweils zwei Dozenten berücksichtigt wurden.

Tabelle 2.4: Altersstruktur aller Dozenten

Alter	**20–29**	**30–39**	**40–49**	**50+**	**gesamt**
absoluter Anteil	1	15	10	8	34
relativer Anteil	2,94%	44,12%	29,41%	23,53%	100%

Aus der Tabelle geht hervor, dass der größte Teil der Dozenten (44,12%) im Altersbereich zwischen 30 und 39 Jahren liegt. Nimmt man noch den einzigen 20–29-jährigen Dozenten mit in diese Klasse auf, so sind es 47,06%. Darauf folgt die Gruppe der 40–49-jährigen mit einem Anteil von 29,41% und zum Schluss die über 50-jährigen mit einem Anteil von 23,53%. Dies deutet darauf hin, dass an den Universitäten mehr jüngere Dozenten IT-Sicherheitsvorlesungen übernehmen. Im nächsten Schritt soll analysiert werden, ob es Unterschiede zwischen den deutschen und den internationalen Universitäten

gibt. Dabei ist allerdings zu berücksichtigen, dass nur eine geringe Anzahl an internationalen Universitäten zur Verfügung steht und deshalb die Betrachtung zwar eine Tendenz, aber keine signifikanten Rückschlüsse erlaubt.

Bei den deutschen Universitäten stellt sich die in Tabelle 2.5 angegebene Situation dar. Der Tabelle ist zu entnehmen, dass an den untersuchten deutschen Universitäten 65% der Dozenten für IT-Sicherheit zwischen 20–39 Jahre alt sind. Die restlichen 35% der Dozenten teilen sich fast gleichmäßig auf die beiden Altersgruppen 40–49 und 50+ auf.

Tabelle 2.5: Altersstruktur der deutschen Dozenten

Alter	**20–29**	**30–39**	**40–49**	**50+**	**gesamt**
absoluter Anteil	1	12	4	3	20
relativer Anteil	5,0%	60,0%	20,0%	15,0%	100%

An den internationalen Universitäten, vgl. Tabelle 2.6, ergibt sich anderes Bild im Vergleich zu den deutschen Universitäten. Die größte Gruppe (42,86%) hier ist in der Altersklasse von 40–49 Jahren angesiedelt. Der kleinste Anteil entfällt auf die Altersklasse der zwischen 30 und 39-jährigen Dozenten, wenn man außer Acht lässt, dass kein Dozent jünger als 30 Jahre ist. Betrachtet man die Klasse der über 40-jährigen, so lässt sich feststellen, dass 78,57% der Dozenten dieser Gruppe angehören.

Tabelle 2.6: Altersstruktur der internationalen Dozenten

Alter	**20–29**	**30–39**	**40–49**	**50+**	**gesamt**
absoluter Anteil	0	3	6	5	14
relativer Anteil	0%	21,43%	42,86%	35,71%	100%

Zusammenfassend lässt sich festhalten, dass die Dozenten von IT-Sicherheitsvorlesungen an deutschen Universitäten jünger sind als ihre Kollegen an internationalen Universitäten.

Als nächstes soll die Hypothese untersucht werden, ob ein Zusammenhang zwischen der Altersstruktur und den gefundenen Clusterstrukturen existiert. Dazu müssen die bei der Clusteranalyse identifizierten Ausreißer außen vor gelassen werden, da sie in keiner

der Klassen vorkommen. Die nachfolgenden Betrachtungen berücksichtigen also *nicht* die Universitäten von Wien, Dresden und Illinois.

In der Tabelle 2.7 sind die Altersangaben der Dozenten aus der innovativen Klasse dargestellt. Die Aufteilung zeigt, dass bei drei der vier Vorlesungen der Dozent relativ jung (≤39 Jahre) ist und nur ein Professor, nämlich Prof. Dr. Christoph Meinel vom Hasso-Plattner-Institut an der Universität Potsdam, in der vierten Altersgruppe (50+) liegt.

Tabelle 2.7: Altersstruktur des innovativen Clusters

Alter	**20–29**	**30–39**	**40–49**	**50+**	**gesamt**
absoluter Anteil	1	2	0	1	4
relativer Anteil	25%	50%	0%	25%	100%

Bei der ausgewogenen, siehe Tabelle 2.8, und der konservativen Gruppe, siehe Tabelle 2.9, sieht das Bild hingegen ausgeglichener aus, da ungefähr 50% der Dozenten zwischen 30–39 Jahren alt sind und die andere Hälfte sich im Alter von 40–50+ Jahren befindet. Es lassen sich hier also keine genaueren Aussagen treffen. Zu bemerken ist, dass in der konservativen Gruppe der prozentuale Anteil der über 50-jährigen im Vergleich zu den beiden anderen Gruppe am höchsten ist.

Tabelle 2.8: Altersstruktur des ausgewogenen Clusters

Alter	**20–29**	**30–39**	**40–49**	**50+**	**gesamt**
absoluter Anteil	0	9	7	4	20
relativer Anteil	0%	45%	35%	20%	100%

Abschließend kann man festhalten, dass es keinen signifikanten Zusammenhang zwischen dem Alter des Dozenten und den von ihm gelehrten Themen gibt. Es lässt sich lediglich die Tendenz erkennen, dass jüngere Dozenten aktuellere Themen lehren, vgl. Tabelle 2.7.

Tabelle 2.9: Altersstruktur des konservativen Clusters

Alter	**20–29**	**30–39**	**40–49**	**50+**	**gesamt**
absoluter Anteil	0	3	1	2	6
relativer Anteil	0%	50%	16,66%	33,33%	100%

2.3.6 Frequenz der Veranstaltungen

Neben der Untersuchung des Alters lassen sich die Cluster noch auf weitere Aspekte hin analysieren. In diesem Abschnitt wird der Frage nachgegangen, wie oft die Vorlesungen der einzelnen Cluster angeboten wurden und ob man daraus Zusammenhänge zu den gelehrten Themen erkennen kann.

Um diese Frage zu untersuchen wurde die in Tabelle 2.10 abgebildete Übersicht erstellt, die auf den recherchierten Daten beruht. Darin sind zu jedem Cluster die zugehörigen Häufigkeiten, wie oft eine Vorlesung stattfand, und der sich aus den Originalwerten ergebende Mittelwert angegeben. Die Mittelwerte resultieren nicht aus den in der Tabelle angegebenen Werten. Deshalb entspricht der Mittelwert in der Zeile „gesamt“ nicht dem Wert, den man bei der Berechnung der drei Mittelwerte dividiert durch drei erhalten würde. Hingewiesen werden muss außerdem auf die Tatsache, dass die Werte nicht der Häufigkeit, wie oft die Vorlesung bisher angeboten wurde, entsprechen, sondern angeben, wie oft der befragte Dozent die Vorlesung gehalten hat. Die Vorlesung selber kann schon häufiger stattgefunden haben. Diese Einschränkung war nötig, da in den meisten Fällen keine ausreichenden Informationen über die Gesamtzahl der stattgefundenen Vorlesungen vorlagen und auch der überwiegende Anteil der befragten Dozenten keine Antwort auf diese Frage gegeben hatte.

Tabelle 2.10: Vergleich der Häufigkeit, wie oft eine Vorlesung stattfand

	Häufigkeit				Mittelwert
	1–3	4–6	7–9	≥ 10	
Innovatives Cluster	3	1	0	0	2
Ausgewogenes Cluster	9	6	1	4	5,7
Konservatives Cluster	2	1	0	2	6
gesamt	14	8	1	6	5,2

Im Folgenden wird Tabelle 2.10 vorgestellt und interpretiert. Wie man erkennt, befinden sich 14 der 29 Vorlesungen, also fast 50%, in dem Intervall mit einer Häufigkeit von 1–3. Im Vergleich zu der Anzahl der Vorlesungen, die mehr als sechsmal stattfanden, kann man feststellen, dass der Großteil der analysierten IT-Sicherheitsvorlesungen wahrscheinlich erst in den letzten Jahren ins Curriculum aufgenommen wurde. Zum Zeitpunkt der Untersuchung (März 2007) hat eine Vorlesung im Mittel 5,2 mal stattgefunden.

Anschließend an die dargestellte globale Betrachtung folgt nun die Auswertung der Häufigkeiten für jede einzelne Klasse. Im innovativen Cluster haben – mit einer Ausnahme – alle Vorlesungen bisher maximal dreimal stattgefunden. Dies deutet darauf hin, dass Themengebiete wie Software- oder Netzwerksicherheit erst seit wenigen Jahren in IT-Sicherheitsvorlesungen gelehrt werden. Im Durchschnitt fand eine Veranstaltung dieses Clusters zweimal statt.

Das ausgewogene Cluster spiegelt ein ähnliches Bild wieder wie das innovative Cluster. Ungefähr 50%, der in diesem Cluster beinhalteten Vorlesungen liegen im Intervall 1–3 und ca. 30% fallen in das Intervall 4–6. Im Unterschied zum innovativen Cluster gibt es Vorlesungen, die mehr als sechsmal angeboten wurden. Durchschnittlich fanden die Vorlesungen dieses Clusters sechs mal statt.

Das konservative Cluster umfasst Vorlesungen in denen hauptsächlich Kryptografie gelehrt wird. Betrachtet man die Häufigkeit, wie oft diese Vorlesungen bisher stattfanden, so stellt sich heraus, dass der Mittelwert im gleichen Bereich wie der des ausgewogenen Clusters angesiedelt ist. Betrachtet man jedes Intervall einzeln, so lassen sich keine besonderen Auffälligkeiten feststellen. Die untersuchten Vorlesungen teilen sich, bis auf eine Ausnahme, gleichmäßig (jeweils 40%) auf das erste Intervall und das letzte Intervall auf. Im Vergleich zu den beiden anderen Clustern ist bei der konservativen Klasse der Anteil von Vorlesungen, die über zehnmal angeboten wurden, am höchsten. Aufgrund der geringen Menge vorliegender Objekte in diesem Cluster ist aber keine allgemeine Aussage möglich.

2.3.7 Verwendete Literatur

Als weiterer Aspekt bei der Untersuchung der Ergebnisse der durchgeführten Clusteranalyse wurde die verwendete Literatur betrachtet. Dazu wurden die erstellten Datensätze der einzelnen Universitäten ausgewertet.

Von den Dozenten der Vorlesungen des innovativen Clusters wurden die folgenden Quellen jeweils zweimal genannt:

- „IT-Crackdown“ von Othmar Kyas und Markus a Campo
- „IT-Sicherheit“ von Claudia Eckert
- „Computer Security“ von Dieter Gollmann
- „Security Engineering“ von Ross Anderson

Man sieht, dass die Bücher „Computer Security“, „IT-Sicherheit“ und „Security Engineering“ benutzt werden, welche zur grundlegenden Literatur im Bereich IT-Sicherheit zählen. Inhaltlich werden in diesen Büchern u. a. grundlegende Themen wie Netzwerkprotokolle und Zugriffskontrolle angesprochen. Interessant ist die Nennung des Buches „IT-Crackdown“, da es in diesem Werk nicht nur um grundlegende Themen zum Schutz des Systems geht, sondern auch Angriffsmethoden und Schwachstellen erklärt werden. Es wird für die IT-Sicherheitsvorlesungen an den Universitäten Potsdam und Dortmund verwendet.

Beim ausgewogenen Cluster verwendeten die Dozenten die folgende Literaturliste, wobei aufgrund der Vielzahl an Quellen nur die angegeben werden, die mindestens viermal genannt wurden. Die Zahl der Nennungen ist in Klammern angegeben.

- „Computer Security: Art and Science“ von Matt Bishop (9)
- „Security in Computing“ von Charles P. Pfleeger und Shari L. Pfleeger (7)
- „IT-Sicherheit“ von Claudia Eckert (7)
- „Security Engineering“ von Ross Anderson (6)

- „Paper und Webquellen“ (6)
- „Cryptography and Network Security“ von William Stallings (5)
- „Applied Cryptography“ von Bruce Schneier (5)
- „Secret and Lies“ von Bruce Schneier (5)
- „Network Security“ von Charlie Kaufman, Radia Perlman, Mike Speciner (5)
- „Computer Security“ von Dieter Gollmann (4)
- „Practical Unix & Internet Security“ von Simson Garfinkel, Gene Spafford und Alan Schwartz (4)
- „Network Security Essentials“ von William Stallings (4)

Diese Liste ist sehr ausgeglichen hinsichtlich der behandelten Themen. Einerseits basieren die Vorlesungen auf Büchern über Kryptografie, wie „Cryptography and Network Security“ von Stallings und „Applied Cryptography“ von Schneier und andererseits greifen viele Dozenten auf Paper und Webquellen zurück. Zudem verwenden vier Dozenten das Buch „Practical Unix & Internet Security“ von Garfinkel, Spafford und Schwartz, welches neben theoretischen Grundlagen, viele praktische Beispiele sowohl für Verteidigung als auch für Angriffsmethoden enthält.

In den Vorlesungen des konservativen Clusters wird laut dem Ergebnis der Clusteranalyse zum größten Teil Kryptografie gelehrt. Dieser Sachverhalt spiegelt sich auch in der erstellten Literaturliste wider, welche im Folgenden angegeben wird. Die Bücher, die nur bei einer Vorlesung verwendet wurden, sind in dieser Liste nicht enthalten.

- „Netzsicherheit“ von Günther Schäfer (3)
- „Cryptography and Network Security“ von William Stallings (2)
- „Network Security Essentials“ von William Stallings (2)
- „Applied Cryptography“ von Bruce Schneier (2)

Alle in dieser Liste aufgeführten Bücher weisen die Gemeinsamkeit auf, dass sie schwerpunktmäßig den Bereich Kryptografie abdecken. Dies deutet darauf hin, dass bei diesem Cluster ein direkter Zusammenhang zwischen der verwendeten Literatur und dem Ergebnis der Clusteranalyse vorliegt.

Betrachtet man die deutschen getrennt von den restlichen Universitäten, dann ergibt sich die in Abbildung 2.6 gezeigte Übersicht; Abbildung 2.5 zeigt entsprechend die vorlesungsbegleitende Literatur der nicht-deutschen („internationalen“) Universitäten (die Zahlen bezeichnen jeweils die Anzahl der Nennungen[2]).

Auffallend bei der Betrachtung der Literatur der internationalen IT-Sicherheitsvorlesungen ist, dass an erster Stelle *Paper und Webquellen* stehen (siehe Abbildung 2.5). Unter diesem Begriff werden hier Publikationen und Webseiten im Bereich IT-Sicherheit verstanden. Beispielsweise basiert die Vorlesung „Internet Security“ der Universität Wien hauptsächlich auf eigener Forschung, Whitepapern, Advisories und verschiedenen Webquellen. Durch die Verwendung der eigenen Forschungsergebnisse und Paper bleibt die Vorlesung auf einem sehr aktuellen Stand und kann sich auch mit neu auftretenden Sicherheitsproblemen auseinander setzen. Am zweithäufigsten wurden die Bücher „Computer Security“ von Bishop und „Security in Computing“ von Pfleeger und Pfleeger genannt, welche beide einen umfassenden Überblick über Themengebiete der IT-Sicherheit geben.

Bei der Analyse der in Deutschland verwendeten Literatur wurde das Buch „IT-Sicherheit“ von Eckert neunmal und damit am häufigsten angegeben (siehe Abbildung 2.6). Dieses Buch ist der Klassiker unter den deutschsprachigen Büchern, die sich mit dem Thema Informationssicherheit in der heutigen Zeit befassen. Weiterhin kann man feststellen, dass sich, im Unterschied zu der bei internationalen Vorlesungen verwendeten Literatur, viele Bücher über Kryptografie in der Liste befinden. Dieser Umstand könnte ein Hinweis dafür sein, dass an den untersuchten deutschen Universitäten größerer Wert auf die Lehre kryptografischer Verfahren gelegt wird.

[2]Da ein Dozent mehrere Bücher angeben konnte, stimmt die Summe der Werte nicht mit der Anzahl der untersuchten Vorlesungen überein.

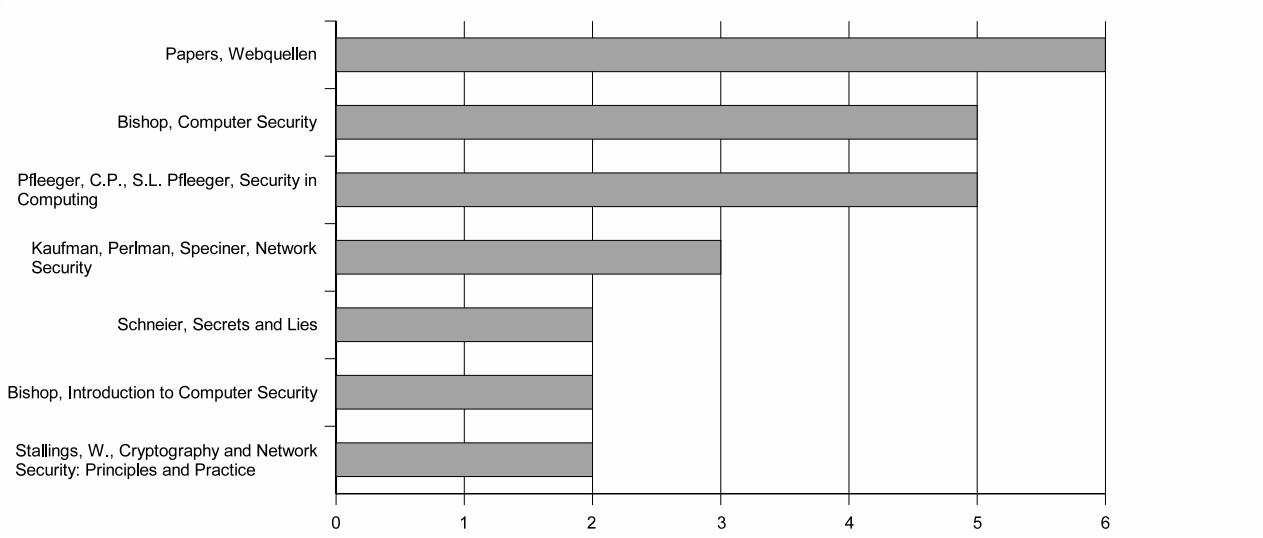


Abbildung 2.5: Vorlesungsbegleitende Literatur (international)

2.4 Studienpläne

Im folgenden werden zuerst existierende Empfehlungen für Studienpläne für Informatik- und IT-Sicherheitsausbildung vorgestellt und danach exemplarisch einige konkrete Vorschläge für bzw. Realisierungen von Studienpläne für IT-Sicherheit skizziert.

2.4.1 Studienplanempfehlungen

Bereits seit den 1960er Jahren existieren Studienplanempfehlungen für den Informatikunterricht, u.a. der „Association für Computing Machinery"(ACM). Allerdings wurden in den ACM-Empfehlungen für das Lehrfach „Computer Science" (eins von fünf Lehrfächern) sicherheitsrelevante Themen erst seit der Fassung vom Dezember 2008 in einige Themengebiete aufgenommen bzw. wurden solche von optionalen zu verpflichtenden Themen [ACM, 2008]. In Deutschland veröffentlichte die Gesellschaft für Informatik e.V. (GI) im Oktober 2006 ihre Empfehlungen zur Berücksichtigung der IT-Sicherheit in der schulischen und akademischen Ausbildung [Gesellschaft für Informatik e.V., 2006]. Hier wird u.a. die Einführung eines verpflichtenden Moduls „Einführung in die IT-Si-

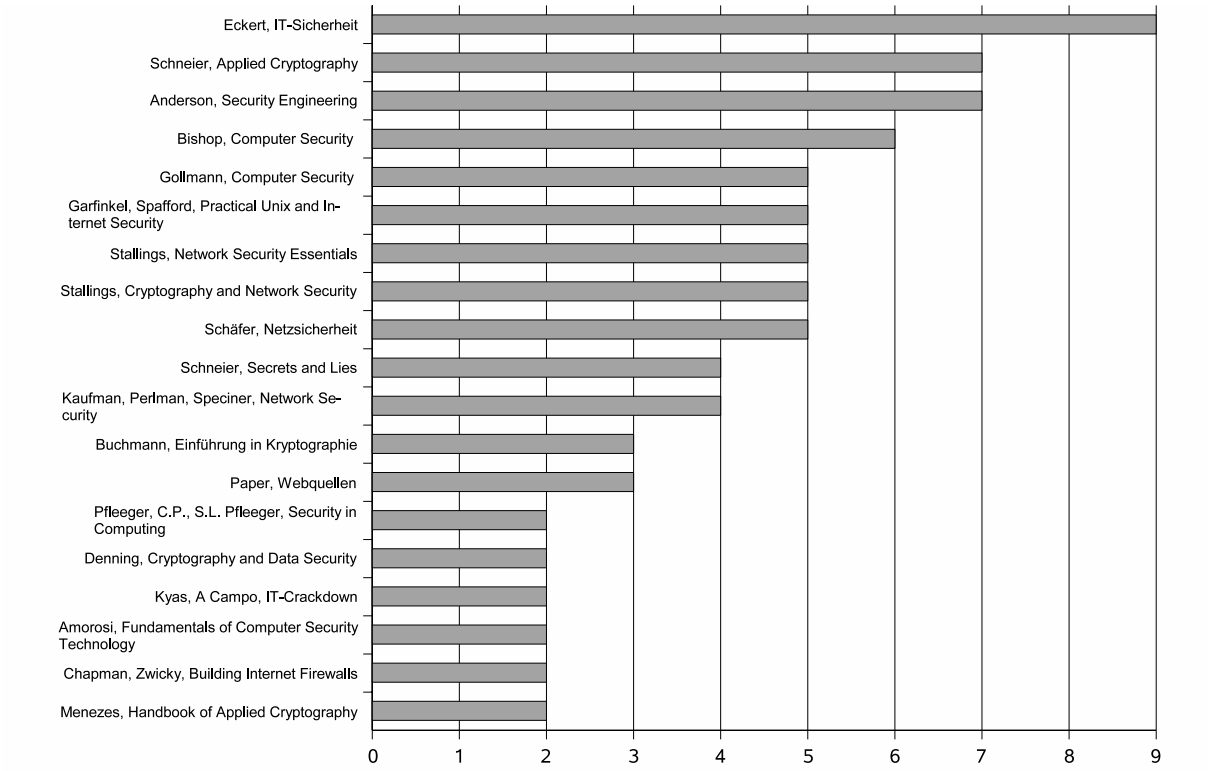


Abbildung 2.6: Vorlesungsbegleitende Literatur (Deutschland)

cherheit“ im Bachelor-Studiengang Informatik gefordert. Darin sollen die grundlegenden Gebiete der IT-Sicherheit eingeführt und Verknüpfungen zu den anderen Basismodulen des Curriculums hergestellt werden. In dem Curriculum des Master-Studiengangs sind weiterführende Veranstaltungen eingeplant, um die erworbenen Kenntnisse vertiefen zu können. Der Schwachpunkt dieser Empfehlung besteht jedoch darin, dass zwar die zu behandelnden Themen angegeben werden, aber „die inhaltliche Tiefe, die genaue Ausgestaltung sowie die Art der Vermittlung“ [Gesellschaft für Informatik e.V., 2006] frei wählbar sind. Es fehlt eine genaue Beschreibung, welche Inhalte der IT-Sicherheit gelehrt und wie sie geübt werden sollen. Außerdem besteht das Problem, welche Themen im Ausgleich für die hinzugenommenen Sicherheitsinhalte weggelassen werden.

2.4.2 Vorschläge für Studienpläne

Zu konkreten Vorschlägen für Studienpläne wird hier eine Auswahl vorgestellt.

Bei dem Modell der RWTH Aachen/Universität Mannheim handelt es sich um eine Kombination von einführender Grundlagenvorlesung mit Spezialvorlesungen (Digitale Forensik, Websicherheit, Betriebssysteme), Seminaren und Praktikum [Dornseif u. a., 2005b].

Carlson beschreibt in seinem Artikel „Teaching Computer Security“ [Carlson, 2004], wie man einen Einführungskurs für Computer- und Netzwerksicherheit entwirft. Dazu gibt er Kursblöcke mit folgenden Themen an, welche auf dem Buch „Counter Hack“ von Skoudis [2002] aufbauen: Linux-Grundlagen, Reconnaissance (Informationssammlung), Netzwerkscans und Exploits, Passwort-Cracken und Angriffe auf Webanwendungen, Netzwerkangriffe, Denial-of-Service-Angriffe sowie Kontrolle über ein erobertes System behalten.

Weitere Veröffentlichungen in dem Bereich sind von Wulf [2003], Mateti [2003], Dhillon u. Hentea [2005], Vigna [2003b], Bogolea u. Wijekumar [2004] und Liles u. Kamali [2006].

2.5 Der offensive Lehransatz

Wie die vorangegangene Betrachtung zeigt, gibt es hauptsächlich zwei Ansätze in der IT-Sicherheitsausbildung: einen eher konservativen Ansatz, dessen Schwerpunkt im Bereich Kryptografie liegt, und einen eher innovativen Ansatz, der aktuelle Themen der IT-Sicherheit favorisiert.

Im konservativen Ansatz, der auch gleichzeitig der klassische Ansatz ist, steht die Lehre von Schutz- und Verteidigungsmaßnahmen im Vordergrund. Er vermittelt die Techniken, die dem Schutz eines Systems und der enthaltenen Informationen dienen. Typischerweise wird hier viel Kryptografie gelehrt, außerdem Konzepte wie Firewalls Intrusion Detection Systeme oder Zugriffskontrollen.

Im innovativen Ansatz werden existierende Systeme und Software auf ihre Sicherheit untersucht. Die dabei erkannten Schwachstellen können einerseits zur Verbesserungen dieser Systeme und für den Entwurf von neuen Systemen genutzt werden. Andererseits soll das Wissen über Angriffsmethoden zu einer verbesserten Strategie in der Abwehr von Angriffen beitragen. Der innovativen Ansatz beschreibt die Vermittlung von Techniken, die darauf abzielen, „etwas kaputt zu machen“. Dabei geht es vorwiegend um Angriffe auf die drei Grundpfeiler der IT-Sicherheit: *Verfügbarkeit*, *Vertraulichkeit* und *Integrität*, z.B. durch Denial of Service-Angriffe oder das Mitlesen (Sniffen) von Informationen im Netzwerk. Die Intention des offensiven Lehransatzes liegt darin, die Methoden der Angreifer kennenzulernen. Dieses Wissen soll helfen, sich der Risiken besser bewusst zu sein und potentielle Gefahren besser einschätzen zu können und somit IT-Systeme besser absichern zu können. In der Literatur [Freiling, 2009] wird der innovative Ansatz mit dem Attribut „offensiv“ belegt, der konservative Ansatz mit „defensiv“.

Was jedoch ist eine „offensive“ und was eine „defensive“ Methode? Generell lässt sich sagen, dass eine offensive Methode eine Methode ist, die ein Angreifer anwendet. Damit lässt sich offensiv und defensiv jedoch nicht ausreichend differenzieren: es gibt Angreifermethoden, die ebenfalls von Verteidigern – d.h. Administratoren und Sicherheitsbeauftragten – angewendet werden, um Systeme abzusichern. Beispiele dafür sind Netzwerksniffen (Angreifer: Mitlesen von sensiblen Daten wie Passwörtern; Administrator: Ermitteln von Problemen im Netzwerk) und Passwortraten (Angreifer: Zugang zu

fremden Accounts; Administrator: Ermitteln von schwachen Passwörtern).

Im folgenden wird deswegen erst eine Definition für „offensiv“ gebildet und anschließend werden Lehrformen für die offensive Lehre vorgestellt.

2.5.1 Definitionen

Zur Lösung des Problems, welche Methode offensiv und welche defensiv ist, wird eine Definition benötigt. Diese wird in diesem Abschnitt aufgestellt.

Eine Definition zu bilden wird dadurch erschwert, dass immer eine Intention enthalten ist. Anders ausgedrückt: Was möchte jemand durch die Anwendung einer Methode erreichen? Auch eine offensichtliche Angreifermethode wie Malware lässt sich zu defensiven Zwecken einsetzen, wenn es sich um einen Wurm handelt, der Sicherheitspatches verteilt. Und eine klassische Schutzmaßnahme wie Kryptografie wird eine offensive Methode, wenn ein Angreifer Daten auf einer fremden Festplatte verschlüsselt und den Zugriff erst nach Zahlung eines Lösegelds freigibt. Diese Definition hilft jedoch nicht, wenn es um den Entwurf von IT-Sicherheitskursen geht, denn dabei gibt es keine Intention bei der Anwendung der Methoden – bzw. ist diese, den Teilnehmern Abwehrmaßnahmen zu zeigen.

Um dieses Problem zu umgehen, wurde eine Möglichkeit gesucht, eine Definition aus der Realität abzuleiten. Sprich: welche Methoden werden in der Ausbildung eingesetzt und lassen sich diese als offensiv oder defensiv klassifizieren? Zu diesem Zweck wurde die in Abschnitt 2.3 vorgestellte Untersuchung durchgeführt. Das dort Ermittelte lässt sich als eine aufzählende Definition von offensiv und defensiv verwenden: offensiv sind die Methoden, die überwiegend im innovativen Cluster gelehrt werden (siehe Abschnitt 2.3.4), defensiv die, die überwiegend im konservativen Cluster gelehrt werden (siehe Abschnitt 2.3.2).

2.5.2 Formen offensiver Ausbildung

In diesem Abschnitt werden in der offensiven Lehre eingesetzte Veranstaltungstypen vorgestellt.

Praktische Kurse

An immer mehr deutschen Universitäten werden Praktika zur IT-Sicherheit angeboten, in denen die Studenten Angreifermethoden praktisch ausprobieren und erleben können. Ein solches Praktikum bietet Studenten die Möglichkeit, beide Seiten der IT-Sicherheit kennenzulernen: mit der Anwendung offensiver Techniken die Rolle des Angreifers und mit defensiven Techniken die Rolle des Administrators von Computersystemen. Die Teilnehmer des Praktikums arbeiten in Teams. Jedes Team administriert eine Anzahl von Computern, auf denen unterschiedliche Betriebssysteme laufen (hauptsächlich Linux und MS Windows). Die Computer sind untereinander vernetzt und haben keine oder nur sehr eingeschränkte Verbindung ins Internet, sodass die Praktikumsteilnehmer Angriffe ausprobieren können, ohne dass es zu Auswirkungen auf andere Netzwerke oder deren Computer kommen kann. Durch die Anwendung der Methoden von Angreifern sammeln die Studenten Erfahrung mit Schwachstellen und Angriffspunkten von Software und Netzwerken. Dieses Wissen versetzt sie in die Lage, ihre Systeme besser gegen Angriffe zu schützen.

Als eine der ersten deutschen Universitäten bot die TU Darmstadt im Jahr 1999 ein Praktikum mit offensiven Aspekten an, das seitdem regelmäßig – mit einer Unterbrechung – stattfindet. Weitere deutsche Universitäten sind die RWTH Aachen (von 2004 bis 2006), die Universität Mannheim (seit 2007), die Universität Magdeburg, die Universität Passau und die Universität Hamburg. Insbesondere an US-amerikanischen Unis sind praktische Kurse mit offensiven Inhalten verbreitet; Beispiele dafür sind die UCSB[3] und Georgia Tech[4].

Capture-the-Flag-Wettbewerbe

In Capture-The-Flag-Wettbewerben (CTF) versuchen die teilnehmenden Teams in die Computer der anderen Teams einzudringen, um so genannte *flags* (engl. für „Flaggen“) zu erobern und gleichzeitig den eigenen Server gegen Angriffe zu verteidigen. Das Spielprinzip wurde von den in den USA populären Capture-The-Flag-Spielen übernommen,

[3] `http://www.cs.ucsb.edu/~vigna/teaching.html`
[4] `http://users.ece.gatech.edu/~owen/`

in denen zwei Teams von Personen eine eigene (reale) Flagge gegen die Eroberung durch das andere Team verteidigen und gleichzeitig versuchen, die Flaggen des anderen Teams zu erobern. Im CTF im digitalen Bereich erhält jedes Team eine Serverinstallation (normalerweise in Form eines Images einer virtuellen Maschine), die vom Veranstalter vorbereitetet und präpariert wurde. Meist handelt es sich um spezielle, vom Veranstalter geschriebene Dienste. Somit ist es nicht möglich, im Internet nach existierenden Schwachstellen von bekannten Diensten und Exploits dafür zu suchen. Die Teammitglieder müssen mit Techniken wie Quellcodeanalyse, Reverse Engineering oder Netzwerkanalyse die Schwachstellen der vorhandenen Dienste und Software erkennen und zum einen herausfinden, wie sich die Schwachstelle beheben lässt, zum anderen Möglichkeiten finden, wie diese ausgenutzt werden können. Auf diese Weise erleben die Teilnehmer sowohl die Rolle des Angreifers als auch die des Verteidigers und benötigen daher Kenntnisse in beiden Bereichen. Erst kurz vor Veranstaltungsbeginn (meist 1 bis 2 Stunden) erhalten die Teams Zugriff auf die Serverinstallation und nutzen diese Zeit, um das System sowie die vorhandene Software und Dienste zu untersuchen. Den erfolgreichen Einbruch in ein System durch das Ausnutzen einer Lücke weist ein Team durch das Einreichen einer Flagge nach, eine nicht erratbare Zeichenfolge, aus der der ausgenutzte Dienst und das davon betroffene Team hervorgeht. Die Flaggen sind nur eine bestimme Zeit gültig (meist im Minutenbereich) und werden in regelmäßigen Abständen von einem so genannten Gameserver auf den Servern der Teams abgelegt. Bei jedem betroffenen Dienst kann die Flagge an einer anderen Stelle bzw. auf eine andere Weise abgelegt werden (z.B. als Eintrag in einem Forum oder in einer Datei). Aufgabe der Teams ist es, dies herauszufinden, um die Flaggen auf einem gegnerischen System finden zu können. Ein Wettbewerb dauert typischerweise zwischen sechs und neun Stunden. Die Teams sind mittels VPN über das Internet verbunden, damit der Wettbewerb in einem geschlossenen Netzwerk abläuft und ohne den Rest des Internets zu beinträchtigen. Für eroberte (gültige) Flaggen erhalten die Teams Offensivpunkte (als Beweis für einen Einbruch); dafür, dass ein Dienst erreichbar ist erhält das jeweilige Team Defensivpunkte (für das Anbieten von (angreifbaren) Diensten). Um zu verhindern, dass Teams alle Dienste des eigenen Servers deaktivieren, um nicht angreifbar zu sein, aber trotzdem andere Server

angreifen, gibt es häufig nur dann Offensivpunkte, wenn der jeweilige auf einem fremden Server ausgenutzte Dienst auf dem eigenen Server aktiv ist. Dadurch sind die Teams auch gezwungen, die eigenen Dienste zu patchen. Durch die Beschreibung von gefundenen Sicherheitslücken und den dafür nötigen Patches können die Teams zusätzlich Punkte erhalten

Es existieren CTF-Wettbewerbe verschiedener Veranstalter und Ausrichtung. Die Wettbewerbe unterscheiden sich in der maximalen Teamgröße, ob die Teams verteilt sind oder lokal teilnehmen, ob Offensive und Defensive nötig ist oder nur Offensive oder ob eine Hintergrundgeschichte existiert. Im folgenden werden einige davon kurz vorgestellt.

An der University of California at Santa Barbara (UCSB) bot Vigna zuerst einen Wettbewerb für Teilnehmer seiner Kurse an. Im Jahr 2003 wurde dieser für Teams anderer US-amerikanischer Universitäten geöffnet und ab dem Jahr 2004 als iCTF für Teams von Hochschulen aller Länder [iCTF]. Im Jahr 2007 nahmen 35 Teams aus 9 Ländern, im Jahr 2008 waren es 38 Teams. Der Wettbewerb findet einmal im Jahr statt. Ein Team darf aus maximal 22 Personen bestehen. In jedem Wettbewerb gibt es eine Hintergrundgeschichte, um die sich die anzugreifenden Dienste thematisch ranken. Im Laufe der Zeit kam zusätzlich zum CTF noch so genannte „Challenges“ hinzu, nicht direkt sicherheitsbezogene Aufgaben, die zusätzliche Punkte bringen. Beispiele für Challenges sind: klingonischen Text in einen Zahlencode übersetzen oder Gesang aus einem Filmausschnitt in Noten umsetzen um daraus das Lösungswort zu bilden.

CIPHER (Challenges in Informatics: Programming, Hosting and ExploRing) [CIPHER] hat seinen Ursprung an der RWTH Aachen. Der Wettbewerb wird von Pimenidis organisiert und wurde zum ersten Mal – bereits für internationale Teams – im Juli 2005 angeboten. Ähnlich wie beim iCTF gibt es Dienste, die erforscht und ausgenutzt werden müssen, aber es gibt keine so stark ausgearbeitete Hintergrundgeschichte. Die maximale Teamgröße beträgt 5 Personen.

Im Rahmen der Konferenz hack.lu [Hack.lu] findet ein CTF-Wettbewerb statt. Dabei handelt es sich weniger um einen klassischen CTF-Wettbewerb als das Lösen von Aufgaben, die mit Sicherheit zu tun haben.

Auf der Hacker-Konferenz DEF CON [Defcon] wird seit 1996 ein CTF-Wettbewerb

veranstaltet. Im Gegensatz zu den meisten anderen Wettbewerben spielen die Teilnehmer vor Ort. Wegen großer Nachfrage gab es in den letzten Jahren eine Qualifikationsrunde. Etwa 8 bis 10 Teams spielen in der Endrunde, die Teamgröße bewegt sich typischerweise im Bereich von 6 bis 10 Personen.

Es existieren weitere CTF-Wettbewerbe im IT-Sicherheitsbereich, eine Übersicht findet sich z.B. auf [Pimenidis].

CTF-Wettbewerbe eröffnen die Möglichkeit, IT-Sicherheitskenntnisse zu messen. Die Fähigkeit eines Teams, Sicherheitslücken zu finden, diese möglichst effizient auszunutzen und zu stopfen sowie dies schriftlich festzuhalten – in Form von Advisories – drückt sich in Punkten aus, die sich am Ende eines Wettbewerbs in einer Rangliste ausdrücken. Damit wird das Wissen und die Fertigkeit, dieses umzusetzen, in einem Wert ausgedrückt (bzw. zwei Werten, da in den meisten Wettbewerben ein Offensiv- und ein Defensivwert angegeben wird) und es ist eine Vergleichbarkeit gegeben.

Wargames

Für ein Wargame erstellt ein Organisator eine Reihe von Aufgaben, die von den Teilnehmern gelöst werden müssen. Diese sind meist level-basiert und können – je nach Art – im Webbrowser oder auf der Kommandozeile bearbeitet werden. Die Aufgaben orientieren sich häufig an Problemen, die ein Angreifer bei dem Versuch einer Systemkompromittierung typischerweise überwinden muss. So müssen z.B. im HTML-Quellcode versteckte Passwörter oder sonstige Informationen gefunden werden und unbekannte Programme reverse engineered werden. Eine bekannte Wargames-Site war Digital Evolution, die jedoch seit Ende 2006 nicht mehr existiert; einige der Wargames von Digital Evolution sind auf [Intruded] gehostet. Weitere Seiten mit Wargames sind [HTS] und [Starfleet]; eine Übersicht findet sich unter [Neworder].

2.6 Für und Wider der offensiven Ausbildung

Ein häufiger Vorwurf gegen die Ausbildung, die offensive Methoden verwendet, lautet, dass durch die Anwendung von Angreifertechniken Studenten zu Hackern[5] ausgebildet werden.

Schauen wir uns die historische Entwicklung an. Offensive Methoden wurden zuerst in der Praxis angewandt: Administratoren nutzten die Methoden von Angreifern, um die Sicherheit ihrer Netzwerke zu überprüfen, indem sie mit entsprechenden Software-Werkzeugen den Netzwerkverkehr beobachteten oder die Passwörter der Benutzer ihrer Systeme auf schwache Passwörter testeten. Solches Vorgehen war jedoch anfangs stark umstritten. Der Vorwurf lautete, dass mit derartigen Werkzeugen Angriffsmethoden für jedermann zugänglich gemacht werden. So wurde Dan Farmer im Jahr 1995 von seinem damaligen Arbeitgeber gekündigt, weil er die Software SATAN veröffentlichte [Markoff, 1995] mit der Administratoren automatisiert nach Schwachstellen auf vernetzten Computern suchen konnten.

Mittlerweile hat sich die Ansicht gewandelt – die Akzeptanz des offensiven Ansatzes lässt sich u.a. in der Literatur an Beiträgen und Artikeln wie denen von Conti [2005] sowie Arce u. McGraw [2004] erkennen – und der offensive Ansatz ist vor einigen Jahren auch in die universitäre Lehre eingezogen. Häufig wird jedoch das Lehren offensiver Methoden immer noch als falsch kritisiert, weil dadurch die Anzahl der „bösen Hacker“ ansteige und folglich der Sicherheitsstand im Internet nicht erhöht, sondern im Gegenteil erniedrigt werde. Jedoch sprechen auch einige Argumente für das Gegenteil. Jede Sicherheitstechnik kann sowohl für gute als auch für schlechte Zwecke genutzt werden. Die Tendenz hin zu Penetrationstests in Unternehmen zeigt, dass offensive Techniken eingesetzt werden können, um den Grad der Sicherheit einer Organisation zu erhöhen. Dies bedeutet, dass Studenten mit Erfahrung in Angriffstechniken nicht notwendigerweise zu bösartigen *Black Hat* Hackern werden müssen, sondern eher gutartige *White*

[5] Der korrekte Ausdruck hier wäre „Cracker“ [RFC 1392, 1993], jedoch hat sich der Ausdruck Hacker eingebürgert. Die ursprüngliche Bedeutung von Hacker ist: jemand, der etwas sehr gut kann (Problemlöser), auch ein sehr guter Programmierer [RFC 1392, 1993]. Die negative, nämlich kriminelle Abwandlung der Bedeutung entstand erst später (etwa seit dem Film „War Games“)

Hat Hacker werden[6]. Offensive Techniken dürfen jedoch nicht für sich allein gelehrt werden. Genauso wie bei Verteidigungstechniken sollte jede Veranstaltung zu IT-Sicherheit von einer elementaren Diskussion rechtlicher Auswirkungen und Ethik begleitet werden. Zu den rechtlich relevanten Regelungen in Deutschland gehören die Paragraphen des Strafgesetzbuches 202a (Ausspähen von Daten), 202b (Abfangen von Daten), 303a (Datenveränderung) und 303b (Computersabotage) sowie seit August 2007 der umstrittene, so genannte „Hackerparagraph“ 202c. In den Veranstaltungen des Lehrstuhls an der Universität Mannheim erfolgt dies durch eine Besprechung der rechtlichen und ethischen Grundsätze zu Veranstaltungsbeginn. Eine weitergehende Methode ist, die Teilnehmer eine Unterlassungserklärung unterschreiben zu lassen. Dies kann den Teilnehmern offensiver Veranstaltungen die Konsequenzen ihres Handelns klarer vor Augen führen und den Veranstalter rechtlich absichern. An manchen US-amerikanischen Universitäten müssen die Teilnehmer vor Veranstaltungsbeginn einem sog. „Background check“ zustimmen, der auf kriminelle Vergangenheit überprüft.

Giovanni Vigna, Organisator des iCTF, äußerte sich in einem Interview mit Technology Review folgendermaßen zu offensiven Methoden und dem Dual Use Prinzip: „Safe-Konstrukteure kennen sich auch mit Safe-Knacken aus“ [Bolduan, 2009]. Schon im Jahr 1868 forderte Hobbs in „The Construction of Locks“ [Hobbs, 1868] eine öffentliche Diskussion über die Sicherheit von Schlössern. Den Bösen seien die Schwachstellen von Schlössern bekannt und die Guten müssten davon erfahren, um gewarnt zu sein.

Der Trend geht auf jeden Fall in Richtung offensiver Methoden: Laut einer Veröffentlichung auf dem News-Portal slashdot will das US-Verteidigungsministerium ein Cyber Security Team aufbauen und sucht dafür in einem Wettbewerb nach Jugendlichen mit Hackererfahrung [Slashdot, 2009][7]. Nach Spiegel-Informationen plant die Bundeswehr in Deutschland mit der „Abteilung Informations- und Computernetzwerkoperationen“ eine spezielle Einheit für die Abwehr von Cyber-Angriffen aufzubauen [Spiegel Online, 2009].

[6]Die Bezeichnungen wurden von Westernfilmen übernommen, in denen die Bösen schwarze Hüte und die Guten weiße Hüte tragen

[7]siehe dort auch die Diskussion über Hacker-Ausbildung an Universitäten

2.7 Zusammenfassung

IT-Sicherheitslehre ist vielschichtig: Nach dem hier vorgestellten Modell gibt es die Ebenen Awareness, Education und Training. Der Stand der IT-Sicherheitsausbildung an Universitäten in Deutschland und international wurde anhand einer Untersuchung der thematischen Inhalte veranschaulicht. Hierbei hat sich gezeigt, dass die untersuchten Veranstaltungen in drei Gruppen eingeteilt werden können, wobei in der ersten Gruppe fast nur aktuelle Themen, in der zweiten hauptsächlich kryptografische Verfahren und in der dritten eine ausgewogene Mischung gelehrt wird. Prinzipiell lassen sich die Vorlesungen in solche mit Kryptografie und solche ohne Kryptografie unterscheiden.

Diese Unterscheidung ist die Basis für die Untersuchung verschiedener Lehransätze, die in Kapitel 4 vorgenommen wird. Zuvor wird jedoch in Kapitel 3 die empirische Testmethodik vorgestellt.

3 Hintergrund

In diesem Kapitel werden Definitionen und Grundlagen vorgestellt, die wesentlich für das Verständnis der Arbeit sind. Im ersten Teil wird die empirische Forschung, im zweiten Metriken vorgestellt. Da eine detaillierte Beschreibung aller empirischen Methoden den Rahmen dieser Arbeit übersteigen würde, sollen hier nur die notwendigen Grundlagen betrachtet werden. Für ein darüber hinausgehendes Interesse seien dem Leser Bortz u. Döring [2006]; Bühner [2004]; Rost [2005] als weiterführende Literatur empfohlen.

3.1 Empirische Methodik

Der Begriff der empirischen Methodik leitet sich aus den griechischen Worten *methodos* („der Weg zu etwas hin“) und *Empirie* („Erfahrungen“) ab. Die empirische Methodik bezeichnet also das geplante und nachvollziehbare Vorgehen, um durch die systematische Auswertung von Beobachtungen und Erfahrungen zu Erkenntnissen zu gelangen. Vier zentrale Begriffe, die die Basis jeder empirischen Untersuchung bilden und noch ausführlich beschrieben werden, sind:

Experiment: Ein Experiment ist eine empirische Untersuchung, die Kausalaussagen zulässt. Die grundsätzlichen Schritte bei der Durchführung eines Experiments bestehen aus dem Aufstellen einer *Hypothese*, dem Feststellen der relevanten *Variablen* und der Entwicklung der Versuchsanordnung mit einem abschließendem Vergleich der gemessenen Ergebnisse (siehe Abschnitt 3.1.1).

Hypothesen: Bilden die Basis für wissenschaftliche Theorien. Diese basieren auf Vermutungen zur Erklärung konkreter Phänomene aufgrund von Beobachtungen und

Überlegungen. Das Ziel einer empirischen Studie ist es, eine *Hypothese* zu stützen oder zu widerlegen. Dabei kann eine *Hypothese* niemals durch eine einzelne Durchführung der Studie bestätigt werden, sondern muss auch Wiederholungen der Studie standhalten, bevor sie als untermauert angesehen werden kann und damit zu einer Theorie wird (siehe Abschnitt 3.1.2).

Variablen: Vor der Überprüfung einer *Hypothese* muss spezifiziert werden, welche Daten zu diesem Zweck erhoben werden müssen. *Variablen* sind dabei die elementaren Bestandteile eines Experiments, wobei vor allem zwischen unabhängigen und abhängigen Variablen zu unterscheiden ist. Dabei bilden die *unabhängigen Variablen* diejenigen experimentellen Bedingungen, die in einer Untersuchung variiert werden, um deren Auswirkungen auf die *abhängige(n) Variable(n)* zu erfassen (siehe Abschnitt 3.1.3).

Versuchsplanung: Auch als „Forschungsdesign“ bezeichnet bildet der Versuchsplan die Grundlage jeder wissenschaftlichen Untersuchung und hat einen entscheidenden Anteil daran, mit welchen statistischen Verfahren eine *Hypothese* zu überprüfen ist und wie aussagekräftig letztendlich die Ergebnisse einer Studie sind (siehe Abschnitt 3.1.4).

3.1.1 Experiment

Die Durchführung einer Studie läuft unter dem Oberbegriff eines Experiments. Nach einer ersten Definition von Greenwood [Klauer, 2005, S. 27] werden in einem Experiment zwei Faktoren, die unabhängige und die abhängige Variable, in unterschiedlichen und kontrollierten Situationen untersucht und in eine ursächliche Beziehung zueinander gebracht. Ein Nachteil dieser Definition ist aber, dass nur jeweils eine Variable betrachtet wird. Gegenwärtige Experimente bestehen aus der Beobachtung und Prüfung mehrerer Variablen, die einen Einfluss auf das Ergebnis der Untersuchung haben können. Basierend auf den Betrachtungen von Wund [Klauer, 2005, S. 28] definiert Rost ein Experiment als die systematische Beobachtung eines Vorgangs, wie sich „unter Konstanthaltung anderer Bedingungen mindestens eine abhängige Variable ändert, nachdem mindestens

eine unabhängige Variable manipuliert worden ist“ [Klauer, 2005, S. 29 f.]. Es müssen folgende Eigenschaften erfüllt sein [Klauer, 2005, S. 30 ff.]:

Planmäßigkeit: Bei einem Experiment handelt es sich um keinen natürlichen Vorgang. Es bedarf eines Versuchsleiters, der einen Versuch selber herbeiführen und die für die Durchführung relevanten Bedingungen eigenständig beeinflussen kann. In der Literatur wird diese Eigenschaft eines Experiments auch oft mit „willkürlich“ bezeichnet.

Wiederholbarkeit: Für ein Experiment ist erforderlich, dass es nicht bei einer einmaligen Aktion bleibt, sondern die Bedingungen so geschaffen werden, dass sich die Durchführung jederzeit unter denselben Voraussetzungen wiederholen lässt und die Ergebnisse somit bestätigt oder auch widerlegt werden können.

Variierbarkeit: Mindestens eine Bedingung (die unabhängige Variable) muss in einem Experiment manipuliert werden, um die Ergebnisse auf Veränderungen zu überprüfen und damit Rückschlüsse auf die variierte Bedingung zu ermöglichen.

3.1.2 Hypothesen

Hypothesen bezeichnen unbewiesene Annahmen oder Behauptungen, dass ein Sachverhalt oder ein Ereignis eintritt, wenn bestimmte Bedingungen vorliegen. Zur Generierung von Hypothesen gibt es verschiedene Strategien, die in der empirischen Forschung als *explorative* Untersuchungen bezeichnet werden. Explorationen bezeichnen das systematische Sammeln von Informationen über einen zu untersuchenden Gegenstand, wobei vier Arten von Strategien zu unterscheiden sind [Bortz u. Döring, 2006]:

Theoriebasierte Exploration: Die theoriebasierte Exploration beschreibt das sorgfältige Durcharbeiten der Fachliteratur und die Recherche nach ähnlichen Studien oder bestehenden Hypothesen bzw. Theorien. Dabei sollen vor allem die Stärken und Schwächen des untersuchten Materials analysiert werden, bevor eine neue Hypothese aufgestellt wird.

Methodenbasierte Exploration: Auch die verwendeten Methoden, mit denen Hypothesen in einer Untersuchung überprüft wurden, können zur Herleitung neuer Hypothesen verwendet werden. Wenn verschiedene Methoden auf denselben Untersuchungsgegenstand angewendet werden, liefern sie nicht notwendigerweise dieselben Ergebnisse. Durch einen Vergleich dieser Unterschiede können eventuelle Eigenheiten des untersuchten Gegenstands identifiziert werden, was zu neuen theoretischen Konzepten führen kann.

Empirisch quantitative Exploration: Diese Methode zur Entwicklung neuer Theorien und Hypothesen verfolgt die Idee, durch eine besondere Darstellung und Aufbereitung von quantitativen Daten bislang unberücksichtigte Muster in den Ergebnissen zu entdecken. Mittels quantitativer Daten können die Ausmaße bestimmter Eigenschaften gemessen und nach statistischer Auswertung der Ergebnisse zum Beispiel als Häufigkeitsverteilungen, Kreuztabellen oder Diagramme dargestellt werden.

Empirisch qualitative Exploration: Analog zur Verwendung der quantitativen Datenanalyse werden hier qualitative Daten dazu verwendet, um neue Hypothesen abzuleiten. Bei qualitativen Daten sind keine Aussagen über Größenverhältnisse der gemessenen Eigenschaften möglich, sondern lediglich, ob zwei Eigenschaften gleich oder ungleich sind, z.B. beim Geschlecht oder der Herkunft von Personen.

Zudem gelten vier wesentliche Eigenschaften, die für jede wissenschaftliche Hypothese erfüllt sein müssen [Bortz u. Döring, 2006]:

1. Eine Hypothese muss sich auf einen *realen Sachverhalt* beziehen, das heißt, die Hypothese ist zum einen nicht trivial und darf sich zum anderen nicht auf spezielle Personen oder Situationen beziehen. Zudem muss dieser Sachverhalt empirisch untersuchbar, also messbar sein, und er muss eine praktische Bedeutung haben.

2. Eine wissenschaftliche Hypothese ist eine *allgemeingütige* Behauptung, die über den Einzelfall oder ein singuläres Ereignis hinausgeht.

3. Hypothesen müssen, zumindest implizit, die Struktur eines sinnvollen *Konditionalsatzes* enthalten, also eine Wenn-Dann- bzw. Je-Desto-Struktur, z.B. „Wenn eine

Person müde ist, dann kann sie sich schlechter konzentrieren“. Konditionalsätze überprüfen, ob bei einer gegebenen Fragestellung zwischen abhängigen und unabhängigen Variablen unterschieden werden kann.

4. Eine Hypothese muss *falsifizierbar* sein, das heißt, es müssen Ereignisse (Falsifikatoren) möglich sein, welche die Hypothese widerlegen können.

Zu den wichtigsten Aufgaben der empirischen Forschung gehört die Überprüfung der zuvor theoretisch abgeleiteten Hypothesen. Die empirischen Daten, die im Laufe einer Studie gesammelt werden, geben selber keine Auskunft darüber, ob eine Hypothese bestätigt oder widerlegt werden kann, sondern bieten nur eine Entscheidungsgrundlage für oder gegen die Hypothese. Dabei besteht, wie bei jeder Entscheidung, die Möglichkeit, sich falsch zu entscheiden. Daher erfolgt die Prüfung von Hypothesen in der empirischen Methodik in der Regel durch statistische Hypothesentests, die auch als *Signifikanztests* bezeichnet werden und die Wahrscheinlichkeit für eine Fehlentscheidung bezüglich der Gültigkeit einer Hypothese berechnen.

Um Irrtümer bei der Hypothesenprüfung zu vermeiden, werden mit einem Signifikanztest zwei Hypothesen überprüft, die so genannte *Nullhypothese* und die *Alternativhypothese*, welche sich gegenseitig ausschließen. Während die Alternativhypothese im Idealfall der zu untersuchenden Hypothese entspricht und die Konsequenz aus einer eingetretenen Bedingung behauptet, besagt die Nullhypothese, dass genau diese Konsequenz nicht eintritt. Damit die Nullhypothese verworfen werden kann, muss die mit dem *Signifikanztest* berechnete *Irrtumswahrscheinlichkeit* unter einer zuvor festgelegten Grenzwahrscheinlichkeit, dem *Signifikanzniveau* (in der Regel 5%), liegen. Würde man oberhalb dieses Wertes die Nullhypothese zugunsten der Alternativhypothese verwerfen, dann ist die Wahrscheinlichkeit hoch, sich zu irren und ein Stichprobenergebnis damit zu Unrecht auf die gesamte Population zu beziehen.

3.1.3 Variablen

Die Überprüfung von Hypothesen erfordert die Zuordnung zu beobachtbaren Phänomenen. Diese Aufgabe wird mit Hilfe von Variablen erfüllt, wobei es sich in der empirischen

Methodik um Eigenschaften von Menschen oder Objekten handelt, die verschiedene Werte (Ausprägungen) annehmen können. So ist z.B. die Körpergröße von Personen eine Variable, da sie von Person zu Person verschieden sein kann, während Konstanten immer nur eine einzige Ausprägung repräsentieren. Die verschiedenen Arten von Variablen, die wichtiger Bestandteil der Durchführung einer empirischen Studie sind, werden hier kurz vorgestellt.

Abhängige und unabhängige Variablen

In einem Experiment variiert der Forscher einzelne Bedingungen, um die Effekte, die sich daraus ergeben, zu betrachten. Dabei werden die veränderbaren Bedingungen als *unabhängige* Variablen bezeichnet, während die Auswirkungen einer Veränderung an der *abhängigen* Variablen geprüft werden. Diese Veränderung kann mit Instrumenten der Datenerhebung, z.B. mündlichen oder schriftlichen Befragungen, Beobachtungen oder Tests festgestellt werden.

Kontroll- und Störvariablen

Der empirische Nachweis eines Zusammenhangs zwischen abhängigen und unabhängigen Variablen ist kein ausreichender Beleg dafür, dass eine abhängige Variable von der unabhängigen kausal beeinflusst wird, also eine Ursache für die Ausprägungen der abhängigen Variablen ist. Eine kausale Beeinflussung kann auch durch *Messfehler* oder *Störvariablen* entstehen. Als *Störvariablen* werden dabei alle Einflussgrößen auf die Ausprägung der abhängigen Variablen bezeichnet, die nicht als Daten erhoben wurden, entweder weil sie als nicht wichtig angesehen oder bei der Versuchsplanung übersehen wurden. Analog bezeichnen *Kontrollvariablen* die Merkmale, deren Ausprägungen vorsorglich erhoben werden und eventuell bei der Interpretation der Ergebnisse hilfreich sein können. Für die Durchführung einer empirischen Untersuchung reicht es aus, die wichtigsten potentiellen Einflussfaktoren zu identifizieren, da es ohnehin nicht möglich sein wird, sämtliche Einflussfaktoren aufzulisten und zu erfassen.

Der schwierigste Part der empirischen Forschung besteht darin, die ausgewählten Variablen in eine messbare Form zu bringen, also von den Merkmalen an Daten zu gelangen.

Dieser Schritt wird auch als *Operationalisierung* bezeichnet und beinhaltet zum Beispiel die Datenerhebung durch Interviews, Fragebogen, Prüfungen, etc., die in Abschnitt 3.1.5 behandelt werden.

3.1.4 Die Versuchsplanung

Der dritte Schritt im grundlegenden Entwurf eines Experiments besteht im Aufbau des Versuchsplans. Ein Versuchsplan ist ein standardisiertes Schema, welches die Basis jeder wissenschaftlichen Untersuchung bildet und beschreibt, wie eine empirische Fragestellung untersucht werden soll. Die Formalisierung eines Versuchsplans hat den Vorteil, dass unabhängig von der inhaltlichen Konzeption des Versuchs anhand einheitlicher Kriterien Aussagen über die Qualität der Untersuchung getroffen werden können. Wesentliche Bestandteile eines Versuchsplans sind neben dem Aufstellen der experimentellen Bedingungen, also der unabhängigen Variablen, folgende Aspekte:

- die Auswahl der am Versuch beteiligten Gruppen,
- die Zuordnung der Versuchspersonen, also die Teilnehmer der Untersuchung, zu den Gruppen sowie
- die Planung der Untersuchungsabfolge.

Diese Planung ist entscheidend dafür, wie aussagekräftig später die Untersuchungsergebnisse sind. In diesem Zusammenhang spricht man auch von der Validität, der Gültigkeit der Ergebnisse, welche je nach Konzeption des Versuchsplans variieren kann. Dabei wird unterschieden zwischen *interner* und *externer* Validität. *Interne* Validität ist dann erreicht, wenn die Veränderung der abhängigen Variable eindeutig auf die Manipulation der unabhängigen Variablen zurückgeführt werden kann, also neben der Untersuchungshypothese keine besseren Alternativerklärungen existieren. Die Anzahl plausibler Alternativerklärungen nimmt dabei durch zusätzliche Faktoren zu, die einen Einfluss auf das Ergebnis ausüben, zum Beispiel durch eine sinkende Aufmerksamkeit der Probanden oder eine ungenaue bzw. fehlerhafte Messung der Merkmale. *Externe* Validität hingegen

ist dann erreicht, wenn das Ergebnis einer Stichprobe auf andere Personen, Zeitpunkte oder Situationen übertragen und verallgemeinert werden kann. Diese ist eventuell gefährdet, wenn eine Untersuchungsumgebung sich zu sehr von natürlichen Gegebenheiten unterscheidet oder eine ausgewählte Stichprobe nicht repräsentativ genug ist. In den folgenden Abschnitten werden die Schritte zur Formalisierung eines Versuchsplans und ihre Auswirkungen auf die Validität der Ergebnisse genauer beschrieben.

Die Auswahl der am Versuch beteiligten Gruppen

Die Auswahl der Teilnehmer für die Durchführung einer Studie erfolgt in der Regel durch die Zusammenstellung einer repräsentativen Stichprobe aus der zu untersuchenden Gesamtmenge, der Population. Die Untersuchung der gesamten Population, die für das Experiment von Interesse ist, ist selten möglich aufgrund eines zu hohen zeitlichen, personellen oder finanziellen Aufwands. Die Verwendung von Stichproben hingegen erfordert weniger Aufwand und ermöglicht auch eine leichtere Auswertbarkeit der Untersuchungsergebnisse [Bortz u. Döring, 2006]. Letztendlich ist der Wert einer Stichprobenuntersuchung davon abhängig, wie gut die ausgewählte Stichprobe eine zu beschreibende Population repräsentiert, um von einer geringen Anzahl Untersuchungsobjekte auf die gesamten Populationsverhältnisse schließen zu können.

In einer empirischen Untersuchung werden in der Regel mindestens zwei Gruppen benötigt, auf welche die Versuchsteilnehmer verteilt werden: eine *Experimental-* und eine *Kontrollgruppe*. Auf die Teilnehmer der *Experimentalgruppe* wird eine spezielle, zu untersuchende Maßnahme angewendet, z.B. die Verabreichung eines bestimmten Medikaments. Auf die Teilnehmer der *Kontrollgruppe* wird die Maßnahme nicht angewendet – bzw. im Fall des Medikaments ein Placebo verabreicht –, um festzustellen, wie sich die Resultate der beiden Gruppen unterscheiden und welche Wirkung die Maßnahme auf die Ausprägung der abhängigen Variable haben.

Die Zuordnung der Versuchspersonen zu den Gruppen

Der zweite Schritt in der Formulierung eines Versuchsplans ist die Verteilung der Versuchspersonen einer entsprechenden Stichprobe auf die Experimental- und Kontrollgrup-

pen. Abhängig vom Untersuchungsziel kann die Gruppenzugehörigkeit durch das Verhalten der Versuchspersonen gegeben sein, z.B. Raucher und Nichtraucher. Sollte dem nicht so sein, so steht der Versuchsleiter vor der Entscheidung, ob diese Zuteilung zufällig (*experimentell*) oder nach einem speziellen Schema (*quasiexperimentell*) durchgeführt wird.

Quasiexperimentelle und experimentelle Versuchspläne *Quasiexperimentelle* Versuchspläne sind dadurch gekennzeichnet, dass die Ausprägungen der unabhängigen Variablen durch eine bewusste, also nicht zufällige Selektion bestimmter Probanden realisiert werden. Dadurch leiden sie im Allgemeinen an einer geringeren Aussagekraft und damit auch an einer geringeren internen Validität als *experimentelle* Pläne, wie in Tabelle 3.1 zu erkennen ist. Der Grund liegt darin, dass bei einer nicht zufälligen Auswahl der Probanden die Ergebnisse nicht eindeutig auf die alleinige Wirkung der unabhängigen Variablen zurückzuführen sind.

In *experimentellen* Versuchsplänen werden die Gruppen hingegen zufällig zusammengestellt. Dadurch können zwar die durch Störvariablen verursachten Unterschiede zwischen den Gruppen nicht eliminiert, aber dafür statistisch konstant gehalten werden, sodass potentielle Störfaktoren in beiden Gruppen in gleicher Weise wirken. In diesem Zusammenhang spricht man auch von einem *statistischen Fehlerausgleich*, was bedeutet, dass zwar jede Zufallsverteilung auch mit Zufallsfehlern behaftet ist, die Vergleichbarkeit zweier oder mehrerer Gruppen aber innerhalb bestimmter statistischer Fehlergrenzen garantiert werden kann [Bortz u. Döring, 2006].

	Interne Validität	Externe Validität
Quasiexperimentell	−	+/−
Experimentell	+	+/−

Tabelle 3.1: Auswirkung der Zuteilungsmethode auf die Validität

Labor- und Feldexperiment *Laborexperimente* finden in einer speziellen Untersuchungsumgebung statt, was dem Versuchsleiter eines Experiments ermöglicht, den Versuchsablauf genau zu überwachen und potentielle externe Einflüsse wie Raumtemperatur, Lärm

oder Luftfeuchtigkeit zu kontrollieren. Dies erhöht die Wahrscheinlichkeit, dass ein kausaler Zusammenhang zwischen der abhängigen und der unabhängigen Variable vorliegt. Die Vorteile einer hohen internen Validität von *Laborexperimenten* gehen allerdings zu Lasten der externen Validität, da aufgrund des unnatürlichen Charakters dieser Untersuchungsmethode und der künstlich geschaffenen Laborbedingungen die Ergebnisse nur bedingt auf die Realität und die Allgemeinheit übertragen werden können.

Genau entgegengesetzt verhält es sich bei *Feldexperimenten*. Diese finden im natürlichen Umfeld der Versuchsteilnehmer statt, wodurch die Untersuchungssituation im Allgemeinen schlechter zu kontrollieren ist und damit auch mehr Alternativerklärungen zulässt. Allerdings können die Ergebnisse durch die natürliche Untersuchungsumgebung leichter auf eine Zielpopulation verallgemeinert werden. Die externe Validität ist somit höher als bei einem *Laborexperiment*, wie Tabelle 3.2 darstellt.

	Interne Validität	Externe Validität
Feldexperiment	−	+
Laborexperiment	+	−

Tabelle 3.2: Auswirkung der Untersuchungsumgebung auf die Validität [Bortz u. Döring, 2006]

Je nach Auswahl von Gruppenzuordnung und Untersuchungsumgebung ergeben sich dadurch mehrere Kombinationsmöglichkeiten, wie sie in Tabelle 3.3 dargestellt sind. Dabei können je nach Kombination die Nachteile des einen Verfahrens durch die Vorteile einer anderen Methode ausgeglichen werden. So ist wie eben dargestellt der Nachteil eines Feldexperiments die mangelnde Kontrolle von Störfaktoren, deren Auswirkungen aber kombiniert mit einer experimentellen Zuordnung durch die Zufallsverteilung minimiert werden können, womit ein Feldexperiment bei zufälliger Verteilung allgemein auch von einer hohen internen Validität profitiert.

Die Planung der Untersuchungsabfolge

Die Planung der Untersuchungsabfolge betrifft die Festlegung des Versuchsaufbaus, auch als Forschungsdesign bezeichnet, und ist entscheidend für die Wahl der Methoden zur

	Experimentelle Zuordnung	Quasiexperimentelle Zuordnung
Feldexperiment	+ intern − extern	− intern + extern
Laborexperiment	+ intern − extern	− intern − extern

Tabelle 3.3: Validität der Ergebnisse in Abhängigkeit der Versuchsplanung [Bortz u. Döring, 2006]

statistischen Überprüfung der Hypothese (Signifikanztests). Entscheidungsgrundlage für die Auswahl eines entsprechenden Designs ist die Anzahl der Gruppen, in welche die Teilnehmer eingeteilt werden, wann die Ausprägungen der abhängigen Variablen erhoben werden und wie sich die Gruppen voneinander unterscheiden sollen. Eine Auswahl häufig gewählter Untersuchungspläne wird im Folgenden vorgestellt [Bortz u. Döring, 2006].

Hinweis: In den Tabellen stehen die Symbole O, X und R für folgende Sachverhalte:

- O: Messen der Merkmalsausprägung der abhängigen Variable
- X: Durchführung einer bestimmten Maßnahme
- R: Randomisierte Zuordnung bei experimentellen Versuchsplänen

Vortest-Nachtest-Plan Der *Vortest-Nachtest-Plan*, wie er in Tabelle 3.4 dargestellt ist, ist der klassische Aufbau in einer empirischen Untersuchung und vor allem bei kleinen Stichproben die beste Wahl. Gegeben sind zwei Gruppen, eine Experimentalgruppe (EG) und eine Kontrollgruppe (KG). Sowohl vor als auch nach der Durchführung einer Maßnahme werden die Ausprägungen eines zu untersuchenden Merkmals erhoben, z.B. durch eine Prüfung, einen Fragebogen oder ein Interview. Ein Nachteil des Vortest-Nachtest-Plans besteht allerdings darin, dass die Probanden durch den Vortest eventuell für die Versuchsdurchführung sensibilisiert werden und sich dadurch während der Untersuchung anders verhalten als ohne Vortest.

Solomon Viergruppenplan Der *Solomon Viergruppenplan* ist eine Erweiterung des Vortest-Nachtest-Plans mit dem Ziel, die Auswirkungen des Vortests, wie sie im vorherigen Forschungsdesign als Nachteil dargestellt wurden, auf die abhängige Variable

Gruppe	Verteilung	Vortest	Maßnahme	Nachtest
EG	R	O	X	O
KG	R	O		O

Tabelle 3.4: Vortest-Nachtest-Plan [Bortz u. Döring, 2006]. Bedeutung der Symbole: O: Messen der Merkmalsausprägung der abhängigen Variable, X: Durchführung einer bestimmten Maßnahme, R: Randomisierte Zuordnung

abschätzen zu können. Dieser Plan (siehe Tabelle 3.5) beinhaltet dabei vier randomisierte Gruppen, wobei die ersten beiden Gruppen genau dem klassischen *Vortest-Nachtest-Plan* entsprechen. In den anderen beiden Gruppen wird allerdings auf die Durchführung eines Vortests verzichtet. Dieser Versuchsaufbau ist sehr aussagekräftig, eignet sich jedoch wegen der doppelten Anzahl benötigter Gruppen nur bei sehr groß angelegten Studien mit entsprechend großen Stichproben.

Gruppe	Verteilung	Vortest	Maßnahme	Nachtest
EG	R	O	X	O
KG	R	O		O
EG	R		X	O
KG	R			O

Tabelle 3.5: Solomon Viergruppenplan [Bortz u. Döring, 2006]

Mehrfaktorielle Pläne Heutige Experimente verfolgen in der Regel nicht mehr das Ziel, die Auswirkungen einer einzigen unabhängigen Variable, sondern die von mehreren Faktoren zu untersuchen. Den einfachsten Fall bildet der *2×2-faktorielle Versuchsplan*, in welchem zwei unabhängige Variablen mit jeweils zwei Ausprägungsstufen betrachtet werden. Auch hier erfolgt die Verteilung der Versuchspersonen auf die verschiedenen Faktorkombinationen zufällig, wobei jede Zelle eine Experimental- bzw. Kontrollgruppe repräsentiert. Ein formalisierter Versuchsplan mit den beiden unabhängigen Variablen A und B hat demzufolge die in Tabelle 3.6 dargestellte Struktur.

		R	R
		A1	A2
R	B1	EG 1	EG 2
R	B2	KG 1	KG 2

Tabelle 3.6: 2×2-faktorieller Versuchsplan

3.1.5 Testtheorie

Bei einem psychologischen Test handelt es sich nach Lienert u. Raatz [1998] um ein wissenschaftliches, also nach bestimmten Regeln durchgeführtes, Routineverfahren zur Operationalisierung von Merkmalen, um eine quantitative oder qualitative Aussage über den relativen Grad der Merkmalsausprägung zu treffen. Methoden zur Datenerhebung sind zum Beispiel Fragebögen, Interviews oder praktische bzw. schriftliche Prüfungen und fallen in den Bereich der experimentellen Diagnostik. Das Ziel psychologischer Tests ist, je nach Reaktion der Teilnehmer auf Testaufgaben oder Fragen (im Folgenden auch als *Items* bezeichnet) auf Persönlichkeitsmerkmale wie Fähigkeiten, Wissen oder Verhalten zu schließen. Dabei stellen sich drei wesentliche Fragen:

- Wie kann anhand dieser Reaktionen auf die Ausprägung eines Merkmals geschlossen werden?
- Wie hängt das Verhalten der Probanden mit dem zu untersuchenden Merkmal zusammen?
- Wie beeinflusst das Merkmal die Reaktion auf ein Item?

Die Beantwortung dieser Fragen ist Gegenstand der Testtheorie, wobei zwischen zwei Klassen unterschieden wird, der *klassischen* und der *probabilistischen* Testtheorie, die im Folgenden miteinander verglichen werden.

Klassische Testtheorie

Die klassische Testtheorie findet in den meisten Tests Anwendung, da sie zum einen schon sehr lange existiert und daher auf viele Erfahrungswerte zurückgegriffen werden kann,

zum anderen ist die klassische Testtheorie leicht anzuwenden und garantiert eine hohe Erfolgswahrscheinlichkeit. Dabei handelt es sich um eine reine Messtheorie, bei welcher dem Probanden in Bezug auf ein Item genau ein Wert zugeordnet werden kann, aber nicht angibt, mit welcher Wahrscheinlichkeit die Testaufgaben von den Probanden beantwortet werden oder wie eine bestimmte Testleistung zustande kommt. Dieser zugeordnete Wert setzt sich zusammen aus dem wahren Wert für die Merkmalsausprägung und einem weiterem Wert, der sich im Gegensatz zum wahren Wert von Messung zu Messung verändern kann: der Messfehler. Es gibt, wie bereits beschrieben, viele Störfaktoren, die vor einer Untersuchung nicht bekannt oder erfassbar sind und wodurch Messfehler entstehen können. Diese Fehler können zufällig bei Einzelpersonen auftreten, was lediglich die Genauigkeit der Ergebnisse, aber nicht deren Richtung beeinflusst. Oder es handelt sich um systematische Fehler, welche sich auf die gesamte Gruppe auswirken und es dadurch zu einer Verzerrung der Ergebnisse einschließlich ihrer Richtung kommen kann.

Für den Zusammenhang von Messfehler und wahrem Wert gelten für die klassische Testtheorie verschiedene Axiome [Bortz u. Döring, 2006]:

1. Der Messwert X setzt sich zusammen aus dem wahren Wert T der Ausprägung und einem Messfehler E. Es gilt also

   $X = TrueScore + ErrorScore = T + E$

2. Wenn man unendlich oft misst, wird es keine systematischen Abweichungen vom wahren Wert mehr geben, es findet also ein Fehlerausgleich statt.

   Für den Messfehler gilt damit

   $\mu(E) = 0,$

   wobei μ der Mittelwert über alle Messungen ist. Damit entspricht das gemessene Ergebnis dem wahren Wert, also

   $\mu(X) = T$

3. Die Höhe des Messfehlers ist unabhängig von den Ausprägungen des untersuchten Merkmals, d.h. Müdigkeit, Konzentration etc. haben dieselben Auswirkungen auf das Ergebnis bei unterschiedlichen Ausprägungen der unabhängigen Variablen.

4. Die Höhe des Messfehlers ist ebenfalls unabhängig von nicht getesteten Merkmalen wie Störvariablen

5. Messfehler sind bei Testwiederholungen voneinander unabhängig. Wenn eine Testperson zum Beispiel beim ersten Test müde oder unkonzentriert ist, dann ist dieselbe Person bei einer Wiederholung des Tests nicht unbedingt genauso müde oder unkonzentriert.

Probabilistische Testtheorie

Bei der *probabilistischen Testtheorie*, die auch als *Item-Response-Theorie* bezeichnet wird, sind im Gegensatz zur deterministischen, klassischen Testtheorie Aussagen über die Wahrscheinlichkeit oder Häufigkeit für bestimmte Antworten von Probanden möglich. Ein sehr bekanntes, aber auch häufig kritisiertes, probabilistisches Modell ist das *Rasch-Modell*, mit welchem die Wahrscheinlichkeit dafür berechnet werden kann, dass ein Proband eine Aufgabe löst oder nicht. Dabei kann aus dem beobachteten Antwortverhalten der Versuchspersonen auf zwei latente, also nicht direkt beobachtbare Variablen geschlossen werden, nämlich die Schwierigkeit einer Aufgabe und die Fähigkeit einer Person, diese Aufgabe zu lösen. Im Gegensatz zur klassischen Testtheorie die Fähigkeit einer Person unabhängig von der Aufgabenschwierigkeit geschätzt werden. Insgesamt ist die Planung nach der probabilistischen Theorie allerdings sehr aufwändig und erfordert große Stichproben, weswegen in der Mehrzahl der Tests die klassische Testtheorie bevorzugt wird.

3.1.6 Fragebogenkonstruktion

Nach Festlegung der Testtheorie sind die entsprechenden Tests bzw. Fragebögen zur Erhebung der zu untersuchenden Merkmale zu entwickeln. Dabei ist zu unterscheiden zwischen der Art des Tests, also welche Merkmale gemessen werden sollen, und die Gestaltung der Testaufgaben, also wie das entsprechende Merkmal zu erfassen ist.

Für die Gestaltung von Testaufgaben und Fragen gibt es zwei Formate: die gebundene und die ungebundene Aufgabenbeantwortung [Bühner, 2004]. Die gebundenen Aufgabenbeantwortung ist die am häufigsten verwendete Methode zur Konstruktion von

Testaufgaben, mit welcher dem Untersuchungsteilnehmer festgelegte Antwortalternativen vorgegeben werden. Vorteil dieses Formats ist die leichtere und vor allem objektivere Auswertbarkeit der Aufgaben, da eine gegebene Antwort eindeutig als richtig oder falsch klassifiziert und dementsprechend bewertet werden kann. Von einer Verwendung von Tests mit nur zwei Antwortalternativen, z.B. Ja/Nein- oder Richtig/Falsch-Behauptungen, ist allerdings abzuraten, da eine Wahrscheinlichkeit von 50%, die korrekte Antwort zu raten, zu einer sehr geringen Aussagekraft des Testergebnisses führt. Eine bessere Alternative bildet der Multiple-Choice-Test, bei welchem den Versuchsteilnehmern viele Antwortmöglichkeiten vorgegeben werden, von denen eine oder mehrere richtig sind. Der Nachteil liegt allerdings im Aufwand bei der Fragenkonstruktion, da es oft schwer ist, sinnvolle Antwortalternativen zu finden und einen gewissen Schwierigkeitsgrad der Fragen zu garantieren.

Der Schwierigkeitsgrad lässt sich mit dem Schwierigkeitsindex messen. Mit dem Schwierigkeitsindex wird der Prozentsatz einer Stichprobe angegeben, der eine Aufgabe korrekt beantworten konnte. Eine leichte Aufgabe, die von vielen Personen korrekt gelöst wird, hat demzufolge einen Schwierigkeitsindex nahe 1. Der Index kann anhand der folgenden Formel berechnet werden [Bortz u. Döring, 2006, S. 218f]:

$$p_i = \frac{\sum_{m=1}^{n} X_{im}}{k_i \cdot n} \tag{3.1}$$

Dabei entspricht $\sum_{m=1}^{n} X_{im}$ den insgesamt erzielten Punkten aller Teilnehmer für eine Frage i und $k_i \cdot n$ den maximal erreichbaren Punkten für diese Frage, wenn die Anzahl aller Antworten n für die Frage korrekt gewesen wären.

Beispiel: Frage 1 wird von 5 Personen korrekt und von 3 Personen falsch beantwortet (also Anzahl aller Antworten $N = 8$). Eine korrekte Antwort soll mit $k_i = 2$ Punkten belohnt werden. Folglich errechnet sich der Schwierigkeitsindex für diese Frage zu:

$$p_1 = \frac{5 \cdot 2}{2 \cdot 8} = 0,625$$

Im Gegensatz zu dem gebundenen Aufgabenformat steht die ungebundene oder auch

freie Aufgabenbeantwortung, in welcher dem Probanden keine Antwortmöglichkeiten vorgegeben werden, z.B. bei der Interpretation oder Deutung bestimmter Vorlagen. So können die Probanden ihr Wissen besser unter Beweis stellen und einfallsreichere bzw. kreativere Antworten geben als durch das bloße Ankreuzen von Antwortalternativen. Allerdings ist die objektive Auswertung vieler unterschiedlicher Antworten sehr aufwendig und problematisch.

Die Auswahl eines Testverfahrens ist davon abhängig, welche Merkmale erfasst werden sollen. Es gibt Leistungstests, um das Wissen oder die persönliche Entwicklung der Versuchsteilnehmer zu bewerten, und Persönlichkeitstests, in welchen mehr die charakterspezifischen Merkmale und Selbsteinschätzungen der Probanden wie persönliche Interessen, Fähigkeiten und Motivation interessieren.

Leistungstests

Leistungstests messen die kognitive Leistungsfähigkeit einer Testperson und setzen voraus, dass eine abgegebene Leistung, z.B. die Beantwortung eines Wissenstests, nach bestimmten Kriterien eindeutig als richtig oder falsch klassifiziert werden kann. Es liegt somit ein rein objektiver Maßstab zu Grunde, welcher garantiert, dass die Beurteilung der Leistung von Prüfer zu Prüfer nicht variiert. Leistungstest werden in *Schnelligkeitstests* und *Niveautests* unterteilt [Bühner, 2004]. In Niveautests steigt mit jeder Frage bzw. Aufgabe ihr Schwierigkeitsgrad, allerdings unterliegen die Testteilnehmer dabei keiner Zeitbegrenzung bei der Bearbeitung der unterschiedlichen Aufgaben. Die Unterscheidung der Testpersonen erfolgt entsprechend des erreichten Fragenniveaus. Bei Schnelligkeitstests hingegen sind die Testaufgaben wesentlich leichter zu beantworten als bei einem Niveautest, allerdings haben die Teilnehmer für die Beantwortung nur begrenzt Zeit. Eine Differenzierung der Teilnehmerleistung erfolgt hier durch die Anzahl der korrekt beantworteten Aufgaben innerhalb einer festgelegten Zeit. Möglich ist auch eine Kombination von Geschwindigkeits- und Niveautests, in welchen die Probanden etwas mehr, aber dennoch beschränkte Zeit bei der Beantwortung schwierigerer Fragen haben.

Persönlichkeitstest

Bei einem Persönlichkeitstest (im Folgenden auch als Awarenesstest bezeichnet) handelt es sich um einen psychologischen Fragebogen, um Personen bezüglich ihrer Einstellung zu einem bestimmten Betrachtungsgegenstand zu befragen. Dabei spielen objektive Maßstäbe wie bei einem Leistungstest keine Rolle mehr und es existieren auch keine richtigen oder falschen Antworten, sondern die Bewertung erfolgt nach einem rein subjektiv festgelegten Bewertungsmaßstab des Testerstellers.

Items werden in einem Persönlichkeitstest in der Regel als Behauptungen formuliert, in welchen der Versuchsteilnehmer auf einer Ratingskala seine Einstellung gegenüber dieser Behauptung einstufen kann. Eine Ratingskala besteht aus mindestens zwei Antwortalternativen, die eine Rangordnung beinhalten, z.B. „Stimme voll zu“ und „Stimme gar nicht zu“. Dabei ist bei der Formulierung der Behauptungen darauf zu achten, dass keine suggestiv wirkenden Begriffe wie „immer, alle, auch, keiner, niemals . . . “ verwendet werden, da diese den Befragten allein durch die Art der Fragestellung zu einer bestimmten Antwort verleiten können, z.B. „Es macht Ihnen doch sicherlich nichts aus, auch Überstunden zu machen?“

Der Nachteil von Awarenesstests ist aber ihre mangelnde Aussagekraft durch ein mögliches unehrliches Antwortverhalten der Testteilnehmer. Dabei sind zwei Arten von Verfälschungen, die zu einer potentiellen Verzerrung der Ergebnisse führen können, zu unterscheiden: *Simulation* und *Dissimulation* [Bühner, 2004]. Mit Simulation wird das Vortäuschen eines Verhaltens bezeichnet, das eine Person normalerweise nicht hat, mit dem Ziel, besonders gute Testergebnisse zu erzielen und dadurch z.B. besonders positiv auf den Versuchsleiter zu wirken. Im Gegensatz dazu bedeutet Dissimulation das Verschleiern und Verbergen von wahrem Verhalten und Einstellungen, um in einem Test absichtlich schlechte Ergebnisse zu erzielen.

Es gibt zwar statistische Verfahren, um durch spezielle Fragestellungen auf ein unehrliches Antwortverhalten der Versuchspersonen schließen zu können, allerdings ist deren Anwendung sehr komplex und zeitintensiv. Daher sei für ein tiefergehendes Interesse Bortz u. Döring [2006] empfohlen. Eine einfachere und verständlichere Alternative ist, den Testpersonen schon vor Beginn der Untersuchung jegliche Beweggründe für unehr-

liche Antworten zu nehmen, zum Beispiel durch eine Anonymisierung des Fragebogens oder durch explizite Hinweise darauf, dass Antworten und Testergebnisse weder zu Konsequenzen noch zu Belohnungen für die Versuchsteilnehmer führen werden.

Letzter Schritt der Planungsphase eines Tests ist die Festlegung des Skalenniveaus, was eine wesentliche Rolle dabei spielt, wie die erhobenen Daten interpretiert werden können und welchen Informationsgehalt sie haben [Bühner, 2004]:

Nominalskala: Mit der *Nominalskala* werden Objekten derart Werte zugeordnet, dass Objekte mit einer gleichen Merkmalsausprägung auch identische Werte erhalten, z.B. bei wertfreien Ja/Nein-Aussagen sowie Fragen nach Geschlecht oder Herkunft. Die Verwendung einer *Nominalskala* erlaubt daher nur Aussagen, ob zwei Merkmale gleich oder ungleich sind.

Ordinalskala: Bei der *Ordinalskala* handelt es sich um eine Rangskala, durch welche im Gegensatz zur *Nominalskala* auch Aussagen, ob die Ausprägung eines Merkmals größer bzw. kleiner ist als die Ausprägung eines anderen Merkmal, erlaubt sind. Allerdings können noch keine Aussagen getroffen werden, wie groß die Abstände zwischen den gemessenen Merkmalen sind.

Intervallskala: Eine *Intervallskala* kombiniert die Eigenschaften von *Nominal-* und *Ordinalskalen* und weist gleichgroße Skalenabschnitte auf, wodurch sowohl größer/kleiner- als auch gleich/ungleich-Aussagen möglich sind. Wesentlicher Vorteil von intervallskalierten Daten ist allerdings, dass auch Aussagen über die Größenunterschiede zwischen zwei Merkmalen getroffen werden können, zum Beispiel, dass ein Merkmal eines Untersuchungsobjekts genau x Punkte stärker oder schlechter ausgeprägt ist als das eines anderen Objekts.

Verhältnisskala: Die *Verhältnisskala* bildet das höchste Skalenniveau, bei welchem im Gegensatz zur *Intervallskala* auch ein absoluter Nullpunkt existiert, z.B. für die Angabe von Temperaturen in Kelvin (nämlich $-273°$) oder das Lebensalter. Das Verhältnis zweier Zahlen entspricht dem Verhältnis der Merkmalsausprägungen der jeweiligen Objekte und ermöglicht u.a. Aussagen, dass sich ein Merkmal um einen bestimmten Prozentsatz verändert hat.

Abschließend ist es für die Aussagekraft der Testergebnisse erforderlich zu überprüfen, ob ein wissenschaftliches Messverfahren bestimmte Qualitätsmerkmale erfüllt. Diese Überprüfung erfolgt anhand verschiedener Gütekriterien, welche auf Basis der Grundaxiome der klassischen Testtheorie definiert wurden [Bortz u. Döring, 2006].

3.1.7 Gütekriterien psychologischer Tests

Gütekriterien psychologischer Tests beziehen sich im Allgemeinen auf zwei Fragestellungen:

- Wird ein zu untersuchendes Merkmal durch die entsprechende Messmethode in einer angemessenen Qualität erhoben?
- Kann anhand des gemessenen Merkmals eine diagnostische Entscheidung mit hoher Qualität getroffen werden (z.B. Kann aufgrund von Testergebnissen ein akademischer Lehransatz bewertet werden)?

Die Beantwortung dieser Fragen erfolgt mit der Überprüfung der drei Hauptgütekriterien *Objektivität*, *Reliabilität* und *Validität*, welche einem hierarchischen Aufbau unterliegen, das heißt, ein nachfolgendes Gütekriterium kann nur dann erfüllt werden, wenn das vorherige Kriterium ebenfalls erfüllt wurde.

Objektivität: Mit *Objektivität* wird die Unabhängigkeit der Untersuchungsergebnisse bezeichnet, also das verschiedene Versuchsleiter bei der Durchführung des Experiments zu denselben Resultaten kommen. Dies geschieht durch eine standardisierte Durchführung und Auswertung des Tests [Bortz u. Döring, 2006].

Durchführungsobjektivität: Ein Testergebnis darf vom Versuchsleiter nicht beeinflusst werden.

Auswertungsobjektivität: Bei der Bewertung mehrerer Testpersonen müssen für gleiche Antworten unter den Probanden auch identische Testwerte vergeben werden.

Interpretationsobjektivität: Die *Interpretationsobjektivität* ist erfüllt, wenn aus denselben Untersuchungsergebnissen verschiedene Interpreten zu den gleichen Schlussfolgerungen kommen. Die Wahrscheinlichkeit dafür ist hoch, wenn sich die Interpretation eines Testwerts an existierenden Referenzwerten orientiert und nicht an individuellen Deutungen.

Reliabilität: Mit Messung der *Reliabilität* wird der Grad der Messgenauigkeit einer Operationalisierungsmethode angegeben. Ein Test ist immer durch Fehlereinflüsse wie Missverständnisse oder das Raten von Antworten belastet und je kleiner diese Fehleranteile sind, desto höher ist auch die Reliabilität. Als Richtwert sollte ein guter Test mindestens eine *Reliabilität* von 80% erreichen, wobei aber auch die Anzahl der Testitems und die Größe der Stichprobe eine Rolle spielt, da erst ab 100 Testpersonen die Reliabilität sicher berechnet werden kann [Mendoza u. a., 2000]. Zur Messung der *Reliabilität* eines Tests gibt es drei unterschiedliche Verfahren [Bortz u. Döring, 2006]:

Testhalbierung: Die Methode der *Testhalbierung* erfordert keinen großen Mehraufwand. Für jeden Versuchsteilnehmer werden zwei Testwerte berechnet, welche jeweils auf der Hälfte aller Testaufgaben beruhen. Die Reliabilität ergibt sich dabei aus der Korrelation der beiden Testhälften.

Retest: Derselbe Test wird derselben Stichprobe nach einem bestimmten Zeitraum wiederholt vorgelegt. Der Vergleich beider Ergebnisse gibt dabei an, wie viel Prozent des Gesamtunterschiedes auf wahre Merkmalsunterschiede zurückzuführen sind. Diese Methode ist aber sehr zeitaufwändig, da die Probanden erneut kontaktiert und für eine Teilnahme an dem Test motiviert werden müssen, wobei aller Erfahrung nach mit einer geringen Rücklaufquote zu rechnen ist.

Paralleltest: Auch der *Paralleltest* erfordert einen erhöhten zeitlichen Aufwand, da insgesamt zwei Tests, die dasselbe Merkmal messen, erstellt werden und von den Probanden direkt nacheinander zu beantworten sind. Je geringer der Unterschied zwischen den beiden Tests ist, desto weniger Fehlereffekte sind

bei den Ergebnissen anzunehmen.

Validität: Die *Validität* eines Tests ist nicht zu verwechseln mit der Validität eines Untersuchungsdesigns (siehe Abschnitt 3.1.4). Die Testvalidität gibt an, ob ein Test wirklich die Merkmale oder Eigenschaften misst, die letztendlich gemessen werden sollen und gilt als das wichtigste, aber zugleich auch am schwierigsten zu bestimmende Testgütekriterium [Bortz u. Döring, 2006]:

Inhaltsvalidität: Ein Test gilt als *inhaltsvalide*, wenn die Fragen und Aufgaben eines Tests das zu messende Merkmal in seinen wichtigsten Aspekten erfassen. Die *Inhaltsvalidität* ist allerdings nicht numerisch zu berechnen, sondern beruht auf subjektiven Einschätzungen.

Kriteriumsvalidität: Die *Kriteriumsvalidität* ist erfüllt, wenn die Ergebnisse eines Testverfahrens mit einem beobachtbaren Merkmal übereinstimmen oder es vorhersagen, z.B. ob die Ergebnisse eines Assessment-Centers zur Personalauswahl auch den beruflichen Erfolg prognostizieren.

Konstruktvalidität: Ein Test gilt genau dann als *konstruktvalide*, wenn aus einem gemessenen Konstrukt neue Hypothesen ableitbar sind, die ebenfalls empirisch überprüfbar sind.

3.2 Metriken für IT-Sicherheitswissen

In vielen Situationen ist es interessant oder sogar notwendig, IT-Sicherheit zu messen. Insbesondere im Rahmen von *Basel II* [Basel II, 2004], welches zum 1. Januar 2007 in Kraft getreten ist, ist dies für viele Unternehmen zu einer wichtigen Angelegenheit geworden. Die vom Basler Ausschuss für Bankenaufsicht herausgegeben Regeln schreiben vor, dass bei der Vergabe von Krediten und der Berechnung von Kreditzinsen neben dem Markt- und Kreditrisiko auch das operationelle Risiko, also die Gefahr von Verlusten durch das Versagen interner Verfahren, Menschen oder Systeme, betrachtet werden muss. Unternehmen sind also darauf angewiesen, durch die Erhöhung ihrer IT-Sicherheit das operationelle Risiko und damit auch ihre Kreditzinsen zu minimieren. Dazu sind

ebenfalls Verfahren anzuwenden, um den aktuellen Stand der Informationssicherheit im Unternehmen zu analysieren und zu bewerten. Ein Ansatz, um diese Herausforderung anzugehen, besteht in der Verwendung von *Messungen* und *Metriken*, mithilfe derer im folgenden *Sicherheitsmetriken* zur Messung von Informationssicherheit entworfen werden. Für zusätzliche Informationen siehe [Mink u. Nowey, 2008].

In den folgenden beiden Abschnitten werden die Begriffe *Metrik* und *Messung* vorgestellt und abgegrenzt und im Anschluss auf Metriken für IT-Sicherheit eingegangen.

3.2.1 Messungen

Die Begriffe *Metrik* und *Messung* sind differenziert zu betrachten. Zwar ermöglichen sowohl Messungen als auch Metriken allgemein die Bewertung eines bestimmten Betrachtungsgegenstandes, allerdings gibt es zwischen diesen beiden Ansätzen wesentliche Unterschiede [Wälchli, 2002]. Messungen sind dadurch charakterisiert, dass mit ihnen einzelne Merkmale eines Gegenstandes oder Prozesses geprüft werden, indem zu einem bestimmten Zeitpunkt das Auftreten eines Merkmales betrachtet wird. So sind zum Beispiel bei der Betrachtung eines E-Mail-Systems folgende Messungen möglich:

- Anzahl E-Mails, die täglich gesendet oder empfangen werden
- Anzahl geschäftlicher und privater E-Mails pro Tag
- Höhe der täglichen Netzwerkbelastung aufgrund von Dateianhängen
- Anzahl ausgelöster Systemfehlermeldungen durch E-Mails pro Tag

Die Ergebnisse dieser Messungen erlauben allerdings nur eine rein quantitative und objektive Bewertung über das E-Mail-System. Zwar können auch Vergleiche zwischen früheren und aktuellen Werten durchgeführt werden, jedoch ist keine Aussage über die Qualität der Ergebnisse möglich, z.B. ob ein E-Mail-System sicher genug ist.

3.2.2 Metriken

Metriken basieren allgemein auf der Durchführung von objektiven Messungen, welche allerdings um zusätzliche bewertende Informationen durch subjektive Erfahrungen ergänzt werden. Metriken sind also im Gegensatz zu Messungen immer auf eine Art intellektuelle Leistung angewiesen und können nicht ausschließlich mittels technischer Verfahren berechnet werden [Wälchli, 2002].

3.2.3 Sicherheitsmetriken

Bei Sicherheitsmetriken handelt es sich um Instrumente zur Bewertung spezieller Sicherheitsaspekte, wie sie vorwiegend in der Industrie angewendet werden. Die Vorteile ihrer Verwendung liegen vor allem darin, dass ein Unternehmen mit ihnen den Fortschritt bei der Implementierung von Sicherheitsstandards verfolgen kann, was durch den Vergleich der gemessenen Merkmale mit einer vorgegebenen Erwartung des Unternehmens über die Ausprägung dieses Merkmals geschieht. Dadurch sind im Gegensatz zu reinen Messungen auch qualitative Aussagen über bestimmte Sicherheitsaspekte möglich, zum Beispiel wie gut die Virenerkennung eines E-Mail-Systems funktioniert, wie effizient getroffene Sicherheitsmaßnahmen sind oder ob ein E-Mail-System heute sicherer ist als gestern.

Weiter bietet die Verwendung von Sicherheitsmetriken auch eine Art Entscheidungshilfe für ein Unternehmen, da sie erlauben, jeden Teil der Unternehmenssicherheit separat zu überprüfen und die Schwachstellen zu isolieren. So kann gezielt in die Bereiche investiert werden, in welchen ein Mangel festgestellt wurde. Durch Vergleiche mit Ergebnissen aus vorherigen Messungen lassen sich zudem die Veränderungen innerhalb eines Zeitraums beobachten und es kann im Falle einer Verschlechterung der Ergebnisse frühzeitig korrigierend eingegriffen werden, bevor ein ernstzunehmendes Sicherheitsrisiko entsteht. Beispiele für Sicherheitsmetriken in der Informationssicherheit sind:

- Prozentsatz von IT-Systemen, die einer Risikobewertung unterzogen wurden,
- Prozentsatz Angestellter, die ein spezielles Training für ihre Tätigkeit benötigen und erhalten haben,

- Verhältnis zwischen Virenalarmen und tatsächlichen Infektionen im Unternehmen oder

- Veränderungen in der Anzahl entdeckter kritischer Schwachstellen auf Servern seit der letzten Messung.

Die Verwendung von Metriken stößt jedoch an ihre Grenzen, wenn die qualitative Bewertung von IT-Sicherheitswissen betrachtet wird, wie es im Rahmen dieser Arbeit erforderlich sein wird. Dies soll an den folgenden zwei Beispielen verdeutlicht werden:

Klassische Tests: Klassische Testverfahren, um das Wissen einer Person beispielsweise im schulischen oder akademischen Umfeld zu bewerten, sind schriftliche Klausuren oder mündliche Prüfungen. Dabei bildet die vergebene Note die Metrik. Durch diese Bewertung sind Abstufungen zwischen den Ergebnissen möglich, allerdings ist eine Note nicht wohldefiniert, da sie keine Möglichkeit bietet, zwei Schulklassen miteinander zu vergleichen.

Zertifikate: Zertifizierungsprogramme und internationale Weiterbildungsstandards für IT-Sicherheit wie der *Certified Information Systems Security Professional* (CISSP), der *TeleTrusT Information Security Professional* (TISP) [TISP] oder *Security+* richten sich an fortgeschrittene IT-Profis. Während CISSP [CISSP] auf US-amerikanische Gegebenheiten ausgerichtet ist und TISP [TISP] für den europäische Markt konzipiert ist, bietet die internationale und herstellerunabhängige *Security+*-Zertifizierung [CompTIA] ein weltweit vergleichbares Niveau. Bei beiden Zertifizierungen ist eine mehrjährige Berufserfahrung im Bereich IT-Sicherheit vorzuweisen und ein umfangreicher Test (im Multiple-Choice-Format) zu bestehen. Zertifizierungen sind für Unternehmen sehr wichtig und bilden eine Entscheidungsgrundlage, z.B. bei der Auswahl neuer Bewerber. Nachteilig ist aber, dass die Ergebnisse aus diesen Zertifizierungstests ausschließlich „bestanden“ oder „durchgefallen“ lauten. Es liegt also nur eine binäre Metrik vor, die keine Möglichkeit biete, zwei Personen mit CISSP-Zertifizierung miteinander zu vergleichen.

Wie diese Beispiele zeigen, mangelt es den bekannten Prüfungsmethoden an der qualitativen Aussagefähigkeit, ob das IT-Sicherheitswissen einer Person größer oder besser als das einer anderen Person ist und wie Unterschiede im Wissenstand sichtbar gemacht werden können. Für einen Vergleich von Lehransätze für die IT-Sicherheit müssen daher andere Messverfahren mit einer neuen, aussagekräftigeren Metrik gefunden werden. Zu diesem Zweck werden Verfahren aus der pädagogischen und psychologischen Forschung, wie sie zu Beginn dieses Kapitels vorgestellt wurden, zu Hilfe genommen und der Vergleich der beiden Lehransätze in Rahmen einer empirischen Studie durchgeführt.

3.3 Zusammenfassung

Im ersten Teil dieses Kapitels wurde in die empirische Methodik eingeführt. Das Gebiet der empirischen Methodik ist ein sehr komplexes und breit gefächertes Gebiet, in welchem viele Sachverhalte zu berücksichtigen sind. Angefangen bei der Untersuchung eines Sachverhalts gilt es, die entsprechende Hypothese mit ihren Untersuchungsbedingungen, also den unabhängigen und abhängigen Variablen sowie einem geeigneten Forschungsdesign, zu erstellen. Den größten Bereich bildet hierbei die psychologische Diagnostik, was die Auswahl eines entsprechenden Messverfahren sowie die Formulierung geeigneter Testaufgaben und Fragen zur Operationalisierung der erforderlichen Daten umfasst. Um dabei eine hohe Aussagekraft der Testergebnisse zu gewährleisten, muss der Test bestimmte Qualitätsmerkmale wie Objektivität, Reliabilität und Validität erfüllen.

Im zweiten Teil des Kapitels wurden Sicherheitsmetriken vorgestellt, welche es ermöglichen, den aktuellen Zustand der Informationssicherheit, z.B. in einem Unternehmen, zu messen und zu bewerten. Allerdings existieren keine Metriken, um das Sicherheitsverständnis von Personen zu messen und Aussagen über dessen Qualität zu machen, wie es für die durchzuführende Studie erforderlich ist.

Die vorgestellten Techniken werden zur Überprüfung von IT-Sicherheitswissen und -fähigkeiten verwendet werden.

4 Methode

In diesem Kapitel wird die Planung und Durchführung der empirischen Studie vorgestellt, die zur Überprüfung der aufgestellten Behauptung benutzt wird. Um den offensiven Lehransatz zu bewerten werden in der Studie empirisch Daten gesammelt. Besser, als nur den offensiven Ansatz allein zu untersuchen, ist, ihn mit einem anderen Ansatz zu vergleichen und dadurch zu bewerten. Hierfür wurde der klassische, defensive Ansatz gewählt, in dem hauptsächlich Verteidigungsstrategien und -techniken gelehrt werden. Zwei Gruppen von Studenten, eine mit offensiv orientierter, die andere mit defensiv orientierter Ausbildung werden in der empirischen Studie verglichen. Für die Grundlagen der angewandten Konzepte empirischen Forschung siehe Kapitel 3.

Das Kapitel besteht aus drei großen Teilen: der Konzeption der Studie, dem Entwurf der in der Studie verwendeten IT-Sicherheitskurse, sowie der Durchführung der Studie. Abschließend werden die gesammelten Erkenntnisse vorgestellt und das Kapitel zusammengefasst. Die Auswertung der gesammelten Daten erfolgt dann in Kapitel 5.

4.1 Konzeption der Studie

Dieser Abschnitt stellt die Planung und Vorbereitung der empirischen Studie vor. Im Speziellen beschäftigt er sich mit dem Aufstellen der Hypothese (Abschnitt 4.1.1), der Auswahl der Variablen (Abschnitt 4.1.2), dem Erstellen der Versuchsgruppen und der Auswahl des Forschungsdesigns (Abschnitt 4.1.3) sowie dem Entwurf der verwendeten Tests (Abschnitt 4.1.4).

4.1.1 Hypothese

Der erste Schritt in einem Experiment ist das Aufstellen einer Hypothese, basierend auf Vermutungen oder explorativen Untersuchungen. Im vorliegenden Fall entstand die Hypothese aus den in der Lehre gemachten, guten Erfahrungen. Ausgehend von der Behauptung, dass der offensive Lehransatz *besser* ist, ergibt sich für die Studie die folgende Forschungshypothese:

> „Studenten, die eine offensive Ausbildung in IT-Sicherheit erhalten, haben ein besseres Verständnis von IT-Sicherheit als Studenten, die eine defensive Ausbildung erhalten."

Für die Untersuchung ist eine genauere Definition erforderlich, was „besser" in diesem Zusammenhang bedeutet. Im Kontext der IT-Sicherheitsausbildung von Studenten kann „besser" auf mehrere Arten verstanden werden:

- ein besseres Verständnis der Schwachpunkte von Sicherheitssystemen
- ein geringerer Zeitaufwand, um sicherheitsbezogene Aufgaben zu erledigen
- eine kontinuierlichere Laufzeit eines administrierten Systems oder
- eine bessere Fähigkeit, sichere Software zu programmieren.

Bei der aufgestellten Forschungshypothese handelt es sich um eine *Unterschiedshypothese*, da sie überprüft, ob sich zwei Maßnahmen, also der offensive und defensive Kurs, in ihrer Wirkung auf das IT-Sicherheitsverständnis unterscheiden. Diese Klassifizierung ist wichtig für die Überprüfung der Hypothese auf Signifikanz, welche in Abschnitt 5.4 vorgenommen wird. Die gewählte Hypothese entspricht den grundlegenden Hypothesenanforderungen (siehe Abschnitt 3.1.2), nämlich dass sie auf einem realen Sachverhalt basiert, allgemeingültig und falsifizierbar ist sowie zu einem sinnvollen Konditionalsatz umformuliert werden kann, was im Folgenden gezeigt wird.

Realer Sachverhalt: Der aufgestellten Hypothese liegt eine praktische Bedeutung zu Grunde, da mit einer Bestätigung der Hypothese die Lehre von IT-Sicherheit verbessert werden kann. Dieser Sachverhalt ist zum einen empirisch untersuchbar, da das IT-Sicherheitsverständnis im Rahmen der Studie messbar gemacht wird, zum anderen bezieht sich die Hypothese nicht auf außergewöhnliche Personen oder Vorgänge.

Allgemeingültigkeit: Die Aussage soll sich auf *alle* Studenten beziehen, die offensive Techniken vermittelt bekommen und geht damit über den Einzelfall hinaus, daher ist die Hypothese als allgemeingültig einzustufen. Das Gegenteil allgemeingültiger Aussagen sind Existenzaussagen, z.B. dass lediglich Studenten existieren, die durch eine offensive Lehre besser und schneller lernen können.

Falsifizierbarkeit: Die Hypothese ist potentiell falsifizierbar, z.B. dadurch, dass die Teilnehmer des defensiv ausgerichteten Kurses ein höheres Sicherheitsverständnis aufweisen.

Konditionalsatz: Die aufgestellte Hypothese enthält implizit die Form eines sinnvollen Konditionalsatzes und kann in einen solchen umgeformt werden, ohne etwas an der Aussage zu verändern. Die als Konditionalsatz formulierte Hypothese lautet:

> „Wenn eine Gruppe von Studenten eine offensive Schulung in Informationssicherheit erhält und eine zweite Gruppe von Studenten eine defensive Schulung, dann ist das Verständnis von IT-Sicherheit bei den Teilnehmern aus der ersten Gruppe größer, als bei denen aus der zweiten Gruppe.“

4.1.2 Variablen

Im zweiten Schritt der Planung werden die zu verwendenden Variablen identifiziert und festgelegt. Da die Auswirkungen der beiden Lehransätze auf das IT-Sicherheitsverständnis beobachtet und gemessen werden sollen, handelt es sich hierbei um die abhängige Variable des Experiments. Die Kurszuordnung (d.h. offensiv, defensiv) wird

zu diesem Zweck variiert und gehört damit zu den unabhängigen Variablen. Neben dieser unabhängigen Variable wird – wie bereits im Experiment von Jonsson u. Olovsson [1997] (siehe Abschnitt 1.4) – das Vorwissen der Untersuchungsteilnehmer, welches den Kenntnisstand der Teilnehmer bezüglich IT-Sicherheit vor der Kursdurchführung repräsentiert, berücksichtigt. Das Vorwissen wird als Kontrollvariable erhoben, das bedeutet, dass es kein expliziter Bestandteil der Forschungshypothese ist, in welcher vorerst nur die alleinige Auswirkung der Kurszuordnung untersucht wird. Die folgende Gliederung fasst alle ausgewählten Variablen mit ihren Eigenschaften und der zu Grunde liegenden Skala zusammen.

Abhängige Variable: IT-Sicherheitsverständnis

- *Polytome* und *diskrete* Variable: Für die Variable existieren viele Abstufungen in ihrer Ausprägung, aber nur endlich viele innerhalb eines begrenzten Intervalls
- *Intervallskaliert*: Da das Wissen nicht negativ werden kann, existiert ein absoluter Nullpunkt. Die Skala ermöglicht somit auch Aussagen bezüglich einer Rangordnung der Ergebnisse.

Unabhängige Variable: Kurszuordnung der Teilnehmer

- *Dichotome* Variable: Die Variable besitzt genau die zwei Ausprägungen „offensiv“ und „defensiv“
- *Nominalskaliert*: Dichotome Variablen sind immer nominalskaliert, da sie nur Aussagen über Gleichheit oder Ungleichheit der Ausprägungen ermöglichen.

Kontrollvariablen: a) Vorwissensgruppe der Teilnehmer

- *Dichotome* Variable mit den beiden Ausprägungen „wenig Vorwissen“ und „viel Vorwissen“
- *Nominalskaliert*

b) Vorwissen der Teilnehmer

- *Polytome* und *diskrete* Variable

- *Intervallskaliert*

Diese Variable muss neben der Vorwissensgruppe gesondert erhoben werden, um über ein anderes Skalenniveau detailliertere Aussagen über die Ergebnisse des Wissenstests und einen möglichen Zusammenhang zum IT-Sicherheitsverständnis machen zu können, als es bei der nominalskalierten Variable „Vorwissensgruppe“ möglich gewesen wäre.

4.1.3 Die Versuchsplanung

Der dritte und letzte Schritt im grundlegenden Aufbau eines Experiments ist das Aufstellen eines Versuchsplans, welcher maßgeblich zur Interpretation und Aussagekraft der Ergebnisse beiträgt. Dazu gehört die Auswahl der Stichprobe und der am Versuch beteiligten Gruppen, die Zuordnung der Untersuchungsteilnehmer und die Planung der Untersuchungsabfolge, also der Festlegung eines Forschungsdesigns.

Festlegung der Zielgruppe und Auswahl der Stichprobe

Als Zielgruppe der Studie – und damit die zu untersuchende Gesamtpopulation – wurden Studenten von Universitäten gewählt.

Zuordnung der Teilnehmer zu den Gruppen

Die Teilnehmer des Kurses mit dem offensiven Lehransatz bilden die Experimentalgruppe (EG) und die Teilnehmer des defensiv ausgelegten Kurses die Kontrollgruppe (KG). Zusätzlich werden die Teilnehmer dieser beiden Gruppen jeweils bezüglich ihres Vorwissens differenziert. Daraus ergeben sich insgesamt vier Gruppen, wie sie in Tabelle 4.1 dargestellt sind.

	Offensiver Kurs	Defensiver Kurs
Wenig Vorwissen	Experimentalgruppe (EG 1)	Kontrollgruppe (KG 1)
Viel Vorwissen	Experimentalgruppe (EG 2)	Kontrollgruppe (KG 2)

Tabelle 4.1: Die vier Stichprobengruppen für die Durchführung der Studie

Die Interessenten werden zufällig auf die Experimentalgruppe und die Kontrollgruppe verteilt. Außerdem werden die Teilnehmer der Kurse nicht darüber informiert, dass es sich bei der Durchführung der Kurse um eine Studie handelte und es sowohl einen defensiven als auch offensiven Kurs gibt. Der Grund dafür ist, dass sich die Teilnehmer möglichst natürlich und unvoreingenommen im Kurs verhalten sollen und auch um zu verhindern, dass sich die Teilnehmer die Kursart nach dem Inhalt bzw. den eigenen Neigungen aussuchen.

Kontrolle der internen Validität – experimentell oder quasiexperimentell Die Studenten werden nach dem Ergebnis im Wissenstest auf die Vorwissensgruppen verteilt. In allen anderen Merkmalen außer der Kurszuordnung und dem Vorwissen sollen sich die Kurse allerdings nicht mehr unterscheiden, um die Anzahl von Alternativerklärungen durch den Einfluss von Störfaktoren zu minimieren. So werden die Kompaktkurse u.a. in derselben Untersuchungsumgebung sowie zu denselben Wochentagen und Uhrzeiten durchgeführt. Da es aber in der Regel nicht möglich ist, alle möglichen Störfaktoren zu identifizieren, wird zur Konstanthaltung der Unterschiede das Prinzip des *statistischen Fehlerausgleichs* (Randomisierungsprinzip, siehe Abschnitt 4.1.3) angewandt, indem die Teilnehmer zufällig auf die Experimental- und die Kontrollgruppe verteilt werden.

Kontrolle der externen Validität – Feld- oder Laborexperiment Nach der Definition eines Feldexperiments findet es – im Gegensatz zum Laborexperiment – nicht in einem speziellen Untersuchungsraum bzw. Labor statt, sondern im natürlichen Umfeld der Versuchsteilnehmer. Obwohl die Studie in einem speziellen Untersuchungsraum – einem Computerraum der jeweiligen Universität – durchgeführt werden soll, kann trotzdem von der Durchführung eines Feldexperiments ausgegangen werden. Das liegt daran, dass es sich in diesem Fall nicht um eine unnatürliche Umgebung für computerversierte Teilnehmer, insbesondere Informatikstudenten, handelt.

Insgesamt handelte es sich bei dieser Untersuchung um ein echtes Feldexperiment, da die Zuordnung der Teilnehmer zu den Gruppen zufällig erfolgte und der Versuch in einem natürlichen Umfeld der Probanden stattfand. Diese Kombination versprach eine hohe interne und eine hohe externe Validität der Untersuchungsergebnisse (siehe

Abschnitt 3.1.4 und Tabelle 3.3 dort).

Auswahl des Forschungsdesigns

Aus der Darstellung der am Versuch beteiligten Gruppen in Tabelle 4.1 ist das entsprechende Forschungsdesign direkt ersichtlich, denn diese Anordnung entspricht genau dem 2×2-faktoriellen Versuchsplan mit den zwei Variablen *Kurszuordnung* und *Vorwissen* und ihren dichotomen Ausprägungen. Insgesamt ergeben sich dadurch vier mögliche Faktorkombinationen, auf welche die Stichproben verteilt wurden. Kombiniert wird das faktorielle Design mit einem *Vortest-Nachtest-Verfahren* (siehe Tabelle 4.2), das bedeutet, dass sowohl vor als auch nach dem Kurs das IT-Sicherheitsverständnis der Teilnehmer anhand eines Tests bzw. Fragebogens gemessen wird. Einen Nachtest erhalten die Teilnehmer im Anschluss an den Kompaktkurs, um zu überprüfen, ob und wie stark sich das IT-Sicherheitswissen der Studenten durch die Teilnahme am Kompaktkurs verändert hat (durch Vergleich zum Vorwert). Zusätzlich kann ein Test einige Zeit nach Ende der Maßnahme durchgeführt werden, um die langfristigen Veränderung seit Ende des Kurses festzustellen.

	Vortest	Maßnahme	Nachtest
Experimentalgruppe	O	X	O
Kontrollgruppe	O		O

Tabelle 4.2: Schema des Vortest-Nachtest-Plans

Die Wahl des Forschungsdesigns ist entscheidend für die Auswahl des *Signifikanztests* zur statistischen Hypothesenprüfung. Bei einem zweifaktoriellen Versuchsplan wie im vorliegenden Fall sind drei Effekte zu analysieren, nämlich die direkten Auswirkungen auf das IT-Sicherheitsverständnis durch die Kurseinteilung (*Haupteffekt* A) und die Vorwissenseinteilung (*Haupteffekt* B) sowie die *Interaktionseffekte* (A×B). Die Haupteffekte überprüfen die Hypothesen, ob die Auswirkungen der durchgeführten Kurse global – also über die gesamte Zielpopulation – interpretiert werden können. Interaktionseffekte hingegen betrachten die differentiellen Wirkungen zwischen den beiden Faktoren Kurseinteilung und Vorwissen [Bortz u. Döring, 2006].

Diese statistische Auswertungsmethode wird auch als *zweifaktorielle Varianzanalyse* (ANOVA[1]) bezeichnet und im Ergebnisteil 5.4.1 detailliert behandelt.

4.1.4 Test- und Fragebogenkonstruktion

Das Testmodell, welches der Studie zu Grunde liegt und bei der Testkonstruktion und Analyse verwendet wird, orientiert sich an der klassischen Testtheorie, da sie leichter anzuwenden, verbreiteter und auch weniger fehleranfällig ist als die probabilistische Testtheorie.

Um die Auswirkungen der beiden Lehransätze zu vergleichen und damit das IT-Sicherheitsverständnis der Studenten messen zu können, sollen insgesamt drei Tests entworfen werden:

1. Ein Awarenesstest, um die Einstellung der Versuchsteilnehmern zu IT-sicherheitsrelevanten Themen zu bestimmen,

2. ein Wissenstest, um die IT-Sicherheitskenntnisse der Studenten zu messen und

3. ein praktischer Abschlusstest, dessen Ergebnisse das Hauptinstrument zur Überprüfung der Hypothese bilden.

Der Wissens- und der Awarenesstest werden in Form eines Fragebogens durchgeführt. Um missverständliche oder zu leichte bzw. zu schwere Fragen bereits im Vorfeld der Untersuchung auszuschließen, wird der Fragebogen einer nicht am Experiment beteiligten Stichprobe vor der Kursdurchführung als Probedurchlauf vorgelegt. Anhand der Auswertung der Ergebnisse und Verbesserungsvorschlägen können problematische Stellen des Tests identifiziert und korrigiert werden. Dazu zählen unklare Fragestellungen, unvollständige Antwortalternativen, Suggestivfragen sowie Fragen, die von fast allen Befragten identisch beantwortet wurden und deswegen nicht zu einer Unterscheidung der Teilnehmer beitragen. In einer Studienarbeit [Sieber, 2008] wurde ein Fragebogen

[1] Analysis of Variance

mit Awareness- und Wissenstest – basierend auf dem in der Studie verwendeten – entworfen, der ausführlicher evaluiert wurde. Er fand keine Verwendung in einer Studiendurchführung, weil er erst danach erstellt wurde, steht aber für zukünftige Experimente zur Verfügung.

Entwurf des Awarenesstests

In dem Awarenesstest werden die Teilnehmern nach ihrer Einstellung zu und ihrem Verhalten mit sicherheitsrelevanten Themen wie Kryptografie, Sicherheitsupdates oder Passwortsicherheit befragt. Die Operationalisierung des IT-Sicherheitsbewusstseins wird aber aufgrund der in Abschnitt 3.1.6 beschriebenen beschränkten Aussagekraft von Awarenesstests nicht dazu verwendet, um eine Entscheidung über den Ausgang der Studie und die Gültigkeit der Hypothese zu treffen, daher fällt dieses Merkmal nicht in die Kategorie einer Kontrollvariablen. Vielmehr sollen die Ergebnisse dazu verwendet werden, um mögliche weitere Auswirkungen des jeweiligen Lehransatzes auf die Teilnehmer festzustellen und bereits vorhandene Ergebnisse anderer Tests zu untermauern.

Bei dem Ratingformat handelt es sich um Fragen mit mehr als zwei Antwortkategorien, die eine gewisse Rangordnung darstellen, z.B. „sehr gut – gut – weniger gut – schlecht“. Gegenüber einfachen Ja/Nein-Fragen sind Ratingformate informationsreicher und erlauben den Versuchsteilnehmern, sich differenzierter bezüglich einer Behauptung zu äußern. Das Kernproblem bei der Konstruktion von Awarenesstests ist die subjektive Festlegung des Bewertungsmaßstabs und damit die Zuweisung von Punktwerten zu den Antwortalternativen einer Ratingskala bzw. der Ja/Nein-Behauptungen. Sollen für Reaktionen, die ein vermeintlich unsicheres Verhalten signalisieren, negative Punkte vergeben werden? Und welches Verhalten kann eindeutig als unsicher eingestuft werden? So stellte sich nach dem Probedurchlauf heraus, dass z.B. die Frage, wie häufig ein Benutzer seine Passwörter ändert, für eine Bewertung des Sicherheitsbewusstseins nicht geeignet ist. Hat eine Person ein sehr gutes, d.h. entsprechend langes und sicheres Passwort, ist es nicht zwangsläufig als schlecht zu interpretieren, wenn diese Person ihr Passwort selten oder gar nicht ändert.

Gänzlich auf eine negative Bewertung bestimmter Fragen zu verzichten erwies sich

als ebensowenig hilfreich. So wurde anhand der Ergebnisse des Probedurchlaufs festgestellt, dass bei der Verwendung einer unipolaren Ratingskala, also einer Skala, die vom Nullpunkt immer nur in eine Richtung verläuft, eine zu geringe Differenzierung in den Ergebnissen vorzufinden ist.

Entwurf des Wissenstests

Während im Awarenesstest die Bewertung der gegebenen Antworten das Problem ist, ist es beim Wissenstest die Auswahl der Fragen zur Prüfung des IT-Sicherheitsverständnis sowie geeigneter und sinnvoller Antwortalternativen. Da die Lehre von Informationssicherheit ein sehr breites Gebiet umfasst, ist es nicht möglich, mit einem kurzen Test dieses gesamte Spektrum abzufragen, um damit Rückschlüsse auf das IT-Sicherheitswissen zu ziehen. Die Testfragen werden so gewählt, dass sie – bis auf einige Ausnahmen – den inhaltlichen Themen der Kurse entsprechen.

Der Wissenstest ist per Definition ein Leistungstest, da alle Antworten eindeutig und objektiv als richtig oder falsch klassifiziert werden können.

Entwurf des Abschlusstests

Der Abschlusstest bildet das entscheidende Messinstrument dieser Studie, um die beiden Kurse und damit die Lehransätze miteinander zu vergleichen und zu bewerten. Hierbei handelt es sich ebenfalls um einen Leistungstests, der allerdings nicht als Niveautest wie der Wissenstest, sondern als praktischer Geschwindigkeitstest konstruiert ist. Die Kursteilnehmer erhalten ein präpariertes Linux-System, mit der Aufgabe, nach Anzeichen für Kompromittierungen oder Konfigurationsfehlern auf diesem System zu suchen und so viele wie möglich zu identifizieren und zu korrigieren. Abbildung 4.1 zeigt die grundsätzliche Idee dahinter: die Teilnehmer beider Gruppen überführen einen PC von einem unsicheren in einen sicheren Zustand. Anhand bestimmter Kriterien, die entweder direkt gemessen oder später ausgewertet werden können, sollen Rückschlüsse auf das IT-Sicherheitsverständnis gezogen und die Lehransätze bewertet werden. Bei den Kriterien, die direkt gemessen werden können, handelt es sich primär um die Anzahl der Kompromittierungen, die ein Teilnehmer gefunden hat, sowie die dafür benötigte

Zeit. Davon abgeleitete Maßzahlen sind die benötige Zeit zum Finden der ersten Kompromittierung, die vergangene Zeit zwischen zwei gefundenen Kompromittierungen und die Anzahl gefundener Lücken nach einer bestimmten Zeit. Nicht direkt messbar sind die Strategie und das Ergebnis. Ersteres kann noch differenziert werden in die Strategie zum Aufspüren einer Lücke und die globale Strategie, nämlich wie die Teilnehmer bei der Analyse und Korrektur des Systems vorgehen.

Um die Idee, die hinter dem Abschlusstest steht, zu verstehen, ist ein kurzer Rückblick auf Abschnitt 3.2 erforderlich. Sicherheitsmetriken sind eine gute Möglichkeit, um z.B. die Lage der IT-Sicherheit eines Unternehmens oder die Auswirkungen von getroffenen Maßnahmen auf die Unternehmenssicherheit zu bewerten. Es wurde allerdings in Abschnitt 3.2 auch gezeigt, dass keine Sicherheitsmetriken für eine qualitative Bewertung des IT-Sicherheitsverständnisses existieren, weswegen für einen Vergleich der offensiven und defensiven Lehransätze eine neue, aussagekräftige Metrik benötigt wird. Das letzte Modul der Kompaktkurse soll ebendiese Forderung erfüllen und mit Hilfe eines praktischen Abschlusstests, der für den offensiven und defensiven Kurs identisch ausgelegt ist, eine solche Metrik aufstellen.

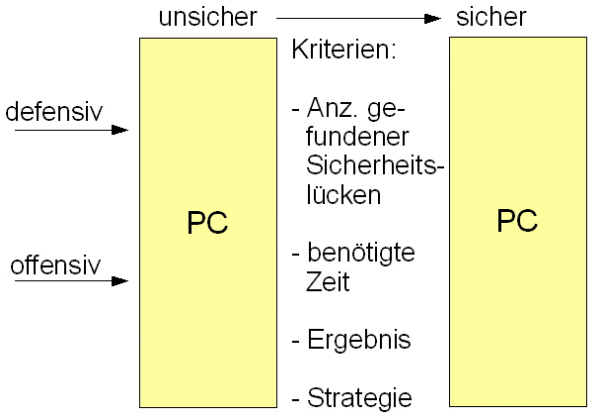


Abbildung 4.1: Konzept des Abschlusstests

Bestimmung der Testgüte

Abschließend zur Testkonstruktion werden die erstellten Tests hinsichtlich der in Abschnitt 3.1.7 beschriebenen Gütekriterien *Objektivität*, *Reliabilität* und *Validität* untersucht.

Objektivität: Die *Objektivität* eines Tests gibt an, ob die Untersuchungsergebnisse vom Versuchsleiter unabhängig sind, sowohl während der Durchführung und der Auswertung, als auch bei der Interpretation von Testergebnissen.

- Die *Durchführungsobjektivität* des Tests ist gewährleistet, da beide Gruppen identische schriftliche Anweisungen auf dem Fragebogen sowie dem Aufgabenblatt des Abschlusstests erhalten. Dadurch ist die Aufgabenstellungen für alle Teilnehmer gleich verständlich geschildert und die Testdurchführung vom Versuchsleiter unbeeinflusst. Theoretisch ist es jedoch möglich, dass allein das Wissen des Dozenten über den Inhalt der Studie (unbewusst) zu einem anderen Verhalten gegenüber den beiden Gruppen geführt hat. Auf diesen so genannten Rosenthal-Effekt wird in Abschnitt 5.5.1 eingegangen.
- Die *Auswertungsobjektivität* wird dadurch garantiert, dass die Antworten jedes Versuchsteilnehmers nach einem festgelegten und standardisierten Bewertungsmaßstab identisch bewertet werden.
- Über die *Interpretationsobjektivität* des Tests kann hier keine Aussage getroffen werden, da bisher keine ähnlichen Untersuchungen durchgeführt wurden, für die Vergleichswerte existieren.

Reliabilität: Die Bewertung der *Reliabilität* eines Tests, also wie viel Fehlereinflüsse in den Testergebnissen enthalten sind, erfolgt durch die Anwendung einer der drei folgenden Methoden: *Paralleltest*, *Testhalbierung* oder *Retest*.

- Für einen vollständigen *Paralleltest*, also der Durchführung von zwei unterschiedlichen Wissenstests für die beiden Kurse, ist eine doppelte Anzahl an Fragen zur Messung des IT-Sicherheitsverständnis notwendig. Allerdings kann

dabei nicht sichergestellt werden, dass die Tests einen identischen Schwierigkeitsgrad aufweisen und die Kursteilnehmer die Aufgaben mit derselben Wahrscheinlichkeit beantworten können.

- Bei Verwendung der *Retest-Methode* wird einige Zeit nach der durchgeführten Maßnahme ein weiterer Test durchgeführt. Das Problem hierbei ist, genug Teilnehmer dazu zu motivieren, mehrere Wochen nach der Kursdurchführung einen weiteren Fragebogen zu beantworten. Eine geringe Rücklaufquote führt zu nicht-aussagekräftigen Ergebnissen.
- Aufgrund der obigen Einschränkungen ist die Messung der Reliabilität mit der *Testhalbierungs-Methode* im vorliegenden Fall die geeignete Wahl. Es ist sinnvoll, den Test nicht nur in zwei Hälften, sondern in so viele kleinste Teile zu zerlegen, wie er Items enthält, also eine Reliabilitätsanalyse über jede einzelne Frage durchzuführen. Das Ergebnis dieser Analyse ist der so genannte *Cronbachs Alphakoeffizient*, welcher mit der folgenden Formel berechnet wird [Bortz u. Döring, 2006]:

$$\alpha = \frac{p}{p-1} \cdot (1 - \frac{\sum_{i=1}^{p} s^2_{Item(i)}}{s^2_{Testwert}}) \tag{4.1}$$

  wobei p die Anzahl der Items eines Tests bezeichnet und s^2 die Varianz der Testwerte. Die Aussagekraft dieser Methode wird jedoch im Fall der durchgeführten Studie gemindert durch die relativ geringe Fragenanzahl des Wissenstests sowie durch eine Teilnehmerzahl, die unter 100 – dem von Mendoza u. a. [2000] geforderten Minimum – liegt.

Validität: Die Bestimmung der *Validität* ist aufwändiger als die Überprüfung von *Objektivität* und *Reliabilität* und kann fast nie als endgültig angesehen werden.

- Die Überprüfung der *Inhaltsvalidität* basiert auf subjektiven Einschätzungen, ob die Testitems das Gebiet der IT-Sicherheit in seinen wichtigsten Aspekten erschöpfend erfassen. Die gesamte Bandbreite der Informationssicherheit lässt sich aufgrund des zu großen Umfangs dieses Wissensgebietes hier nicht

abdecken, daher konzentrieren sich die Fragen weitestgehend auf die Themen, die auch in den Kursen behandelt werden (siehe Abschnitt 4.2). Der Test ist daher in Bezug auf die Kompaktkurse als inhaltsvalide anzusehen, allerdings nicht in Bezug auf den gesamten IT-Sicherheitsbereich.

- Der Test erfüllt die Anforderungen der *Kriteriumsvalidität*, da die Ergebnisse des Wissenstests der Teilnehmer mit viel Vorwissen mit einem beobachtbaren Verhalten, hier der Bearbeitung des Abschlusstests, übereinstimmen. Wie die Ergebnisse der beiden Tests in Abschnitt 5.2 bzw. 5.3 zeigen werden, haben die Teilnehmer mit einem höherem Vorwissen auch den Abschlusstest besser bearbeiten können.
- Die *Konstruktvalidität* ist erfüllt, da aus dem Ergebnis der Studie eine Hypothese über eine differentielle Wirkung von Vorwissen und Kurszuordnung ableitbar ist (siehe Abschnitt 5.4.1), welche in einer weiteren Studie überprüft werden kann.

Nachdem die grundsätzlichen Voraussetzungen für das Experiment geplant und die Regelungen für die Auswahl der Teilnehmer festgelegt wurden, behandelt der folgende Abschnitt die Konzeption der Kurse, also die inhaltliche Planung des Theorieteils und der Übungen.

4.2 Konzeption der Kurse

Für die Studie war es notwendig, zwei Kurse zu konzipieren, einer für die defensiv orientierte Ausbildung, einer für die offensiv orientierte. Die Wahl fiel auf einen dreitägigen Kompakt- bzw. Crashkurs. Ein Kompaktkurs hat im Vergleich zu einer regulären (d.h. semesterlangen) Lehrveranstaltung den Vorteil, dass er häufiger wiederholt werden kann, um mehr Daten für die Studie zu erfassen oder eventuelle Fehler in der Planung zu korrigieren.

Die Kurse bestehen aus neun Modulen, deren Durchführung auf drei aufeinanderfolgende Tage verteilt ist. Bei sieben Modulen handelt es sich um einen Themenblock aus dem Bereich IT-Sicherheit, bei einem um einem um eine Einführung und beim letzten

um einen praktischen Test. Am Anfang jeden Themenblocks werden in einem ca. 30-minütigen Vortrag die theoretischen Konzepte des entsprechenden Themas vorgestellt, gefolgt von einem etwa einstündigen praktischen Übungsteil mit anschließender Diskussion der Lösungen. Der Theorieteil ist für beide Kurse identisch gehalten, da die Teilnehmer identisches Vorwissen erhalten sollen und eine Differenzierung zwischen dem offensiven und dem defensiven Lehransatz in den Übungen stattfinden soll. Als Betriebssystem wird Linux genutzt, da es frei verfügbar ist, eine Vielzahl von Einstellungsmöglichkeiten erlaubt und viele Werkzeuge bereits mitbringt.

Ein grundlegendes Konzept der Kurse ist, dass vorhandene (grafische) Konfigurationswerkzeuge nur eingeschränkt genutzt werden, sondern Einstellungen direkt in Konfigurationsdateien mithilfe eines Texteditors vorgenommen werden. Dieses direkte Arbeiten am System soll zu einem besseren Verständnis führen und ist unabhängig von der verwendeten Linux-Distribution bzw. dem verwendeten System.

Eine Übersicht über die Kursinhalte ist in Tabelle 4.3 dargestellt. Die folgenden Abschnitte stellen die einzelnen Module vor. Nach einer Beschreibung der Inhalte folgt jeweils eine Tabelle, die die Unterschiede im praktischen Teil zwischen dem offensiven und dem defensiven Kurs gegenüberstellt und damit veranschaulicht.

4.2.1 Modul 1: Einführung

Die Einführungsveranstaltung dient dazu, den Teilnehmern die wesentliche Grundlagen für das Verständnis der folgenden Kursinhalte zu vermitteln und eine gemeinsame Basis zu schaffen. Dazu gehört eine Einführung in Linux mit grundlegenden Eigenschaften von Linux, eine Vorstellung der wichtigsten Kommandozeilenbefehle und eine Übersicht über die Verzeichnisstruktur. Neben einer kurzen Zusammenfassung der wichtigsten Netzwerkgrundlagen (u.a. TCP/IP-Protokollfamilie, Ports, Subnets) erfolgt eine Einführung in die Programmierung mit C, da vor allem in den Modulen *Softwaresicherheit* und *Malware* Kenntnisse in dieser Programmiersprache benötigt werden. Zudem werden auch ethische und rechtliche Aspekte von IT-Sicherheit angesprochen, zusammen mit dem eindringlichen Appell, die im Kurs gelernten Methoden nur in isolierten Netzwerken und nur mit dem Einverständnis der Betroffenen anzuwenden.

Tag 1	Tag 2	Tag 3
1. **Einführung**: Ethik Linux-Grundlagen Programmierung in C Netzwerk-Grundlagen	4. **Netzwerksicherheit** 1: Sniffen Port Scanning	7. **Websicherheit**: SQL Injection Webgoat Schwachstellen
2. **Unix-Sicherheit**: Passwortsicherheit Zugriffskontrolle Kompromittierungen	5. **Netzwerksicherheit** 2: Spoofing TCP-Hijacking Denial of Service SSH	8. **Malware**: Viren Würmer Trojaner Rootkits
3. **Softwaresicherheit**: Speicherorganisation Buffer Overflows Format Strings Race Conditions	6. **Firewalls**: Konzept Architektur Konfiguration	**Abschlusstest**

Tabelle 4.3: Überblick der Kursinhalte

Die Übung ist für den offensiven und den defensiven Kurs identisch gestaltet, um den Teilnehmern beider Kurse ein Grundverständnis für die nachfolgenden Themen und Übungen zu vermitteln. Neben der Konfiguration der verwendeten virtuellen Maschine und Einrichtung der Netzwerkverbindung (manuelles Eintragen von IP-Adresse, Netzmaske, Standard-Gateway und Rechnername), erlernen die Teilnehmer den Umgang mit verschiedenen Kommandozeilenbefehlen, dem Linux Standard-Text-Editor *vi* und dem C-Compiler *gcc*.

4.2.2 Modul 2: Unix-Sicherheit

Das Kapitel der Unix-Sicherheit behandelt im ersten Teil allgemeine Sicherheitsaspekte von Passwörtern, also was schwache und starke Passwörter sind, wie die Passwortverschlüsselung unter Linux und Windows funktioniert, sowie typische Angriffe auf Passworte wie Brute-Force- oder Wörterbuchangriffe.

Der zweite Teil des Vortrags stellt die wesentlichen Aspekte der Zugriffskontrolle unter

Linux vor. Dazu gehört die Beschreibung des *SUID-Bits*, den potentiellen Gefahren durch SUID-Root-Dateien und wie diese auf dem System gefunden werden können.

Zudem werden die Möglichkeiten behandelt, um auf dem System nach Anzeichen für Kompromittierungen zu suchen, nämlich durch Überwachung der laufenden Prozesse, ob sie unter ihrer korrekten Benutzerkennung laufen bzw. ungewöhnlich viel Rechenzeit verbrauchen, oder die Überprüfung von Protokolldateien auf unerwartete Neustarts von Diensten und ungewöhnliche Fehlermeldungen.

Der gemeinsame Teil der Übung besteht in der Verwendung des Password Recovery Programms *John the Ripper*, um die Passwörter der auf dem System eingetragenen Benutzer zu überprüfen. Zudem sollen die Teilnehmer nach SUID-Root-Dateien suchen und dabei die Anwendung *cat* als potentielles Sicherheitsrisiko identifizieren, da damit unberechtigte Benutzer Dateien lesen können, für die sie keine Berechtigung haben. Am Ende des Moduls soll jeder Teilnehmer ein sicheres Passwort für seinen Account setzen.

Unix-Sicherheit

Offensiv	Defensiv
Detaillierte Betrachtung der Funktionen von *John the Ripper* inkl. Einbindung zusätzlicher Wörterbücher	Erstellen sicherer *Passwortrichtlinien*
Ausführung systemkritischer Befehle mit dem *sudo*-Befehl	Einrichtung einer *Passwortalterung*
Beseitigen von Spuren	*Suchen* nach Spuren

Tabelle 4.4: Unterschiedliche Übungsinhalte des Moduls Unix-Sicherheit

4.2.3 Modul 3: Softwaresicherheit

Da Fehler in Anwendungs- oder Systemsoftware mit zu den häufigsten Ursachen für Kompromittierungen und Sicherheitsverletzungen eines Systems gehören, befasst sich dieses Modul detailliert mit den Risiken durch unsicher programmierte Software. Dies beinhaltet eine Einführung in die Speicherorganisation von Linux (u.a. Stack, Heap, Rücksprungadresse) und wie ein Angreifer mittels Privilegeskalation durch *Buffer Overflows*, *Race Conditions* oder *Format String-Angriffe* schadhaften Code ausführen bzw.

Rootrechte auf dem System erlangen kann. Zudem werden zu jedem dieser Angriffe auch die entsprechenden Schutzmaßnahmen vorgestellt.

In dem gemeinsamen Übungsteil sollen die Teilnehmer die Funktionsweise von *Buffer Overflows* anhand eines unsicher programmierten Programms austesten.

Softwaresicherheit

Offensiv	Defensiv
Ausführliche Betrachtung von *Buffer Overflows*	Ausnutzen von *Race Conditions* *Quellcodeanalyse* von C-Programme und anschließende *Fehlerkorrektur*

Tabelle 4.5: Unterschiedliche Übungsinhalte des Moduls Softwaresicherheit

4.2.4 Modul 4: Netzwerksicherheit 1

In diesem und dem folgenden Modul werden den Teilnehmern verschiedene Netzwerkangriffe, aber auch Methoden zur Erhöhung der Netzwerksicherheit, vorgestellt. Dabei behandelt der Block *Netzwerksicherheit 1* in der ersten Hälfte die Funktionsweise von *Netzwerk-Sniffern* in geswitchten Netzwerken, um mittels ARP- oder MAC-Spoofing den Datenverkehr in einem Netzwerk aufzuzeichnen und auswerten zu können, sowie die Möglichkeiten, um einen aktiven Sniffer auf dem System ausfindig zu machen. In der zweiten Hälfte dieses Moduls werden die Grundlagen von Netzwerk-Scannern, die unterschiedlichen Scan-Methoden (z.B. *SYN Scanning, Idle Scanning*) und die Funktionen ausgewählter Netzwerk- bzw. Vulnerability-Scanner (z.B. *nmap, Nessus [Tenable Network Security]*) betrachtet.

Der gemeinsame Teil der Übung besteht darin, mit einem Netzwerk-Sniffer den TCP-Verbindungsaufbau zu beobachten und die Passwörtern im Datenstrom zu identifizieren, wenn eine FTP- oder Telnetverbindung zu einem anderen Rechner errichtet wurde. Eine weitere gemeinsame Übung besteht in der Verwendung des Portscanners *nmap* und dem ausgiebigen Testen seiner Eigenschaften, z.B. dem Ermitteln der IP-Adressen aller im Kursnetz vorhandenen Rechner, der Durchführung von *SYN*- oder *XMAS-Scans* und

die Ermittlung des installierten Betriebssystems durch *OS-Fingerprinting*.

Netzwerksicherheit 1

Offensiv	Defensiv
Sniffen von Passwörtern Sniffen mittels ARP-Spoofing Durchführung von Idle-Scanning	Detektieren von Sniffern auf dem System

Tabelle 4.6: Unterschiedliche Übungsinhalte des Moduls Netzwerksicherheit 1

4.2.5 Modul 5: Netzwerksicherheit 2

Im zweiten Teil der *Netzwerksicherheit* werden verschiedene *Spoofing*-Methoden (z.B. *ARP-, IP-, Mail-* oder *DNS-Spoofing*) zur Vortäuschung einer anderen Identität behandelt. Ein weiterer Bestandteil dieses Moduls umfasst die Grundlagen von SSH und der Einrichtung einer Public-Key-Authentifizierung. Im letzten Abschnitt werden Kenntnisse von *Denial of Service-Angriffen*, also Angriffen, die auf die Verfügbarkeit von Informationen und Systemen abzielen, vermittelt.

In der gemeinsamen Übung sollen die Teilnehmer einen *SYN-Flood-Angriff* auf den eigenen Rechner durchführen und dabei die Auswirkungen beobachten, wenn die Option SYN-Cookies, die einen Rechner vor SYN-Flood-Angriffen schützen sollen, aktiviert bzw. deaktiviert ist. In diesem Modul konnte, wie Tabelle 4.7 zeigt, eine sehr gute Abgrenzung von offensiven und defensiven Inhalten erreicht werden.

4.2.6 Modul 6: Firewalls

Bei Firewalls handelt es sich entweder um Sicherheitskomponenten in einem Netzwerk (Hardware-Firewall) oder auf einem Computersystem (Personal Firewall). Ziel von Hardware-Firewalls ist es, den Datenverkehr zwischen Netzsegmenten abzusichern und die interne Struktur eines Netzes zu verbergen, während eine Personal Firewall nur den ein- und ausgehenden Datenverkehr zwischen Rechner und Netz filtert, aber nicht den

Netzwerksicherheit 2

Offensiv	Defensiv
Fälschen von MAC-Adressen (MAC Cloning)	Konfiguration eines SSH-Servers und Clients
Durchführung von ARP-Spoofing	Überprüfen des eigenen Systems mit *Nessus*
Scannen und analysieren des Netzwerks	Konfiguration und Anwendung einer E-Mail-Verschlüsselung
Sniffen von Passwörtern und Webinformationen	

Tabelle 4.7: Unterschiedliche Übungsinhalte des Moduls Netzwerksicherheit 2

zwischen zwei Netzwerken. In diesem Modul werden den Teilnehmern die verschiedenen Firewallkonzepte, Implementierungsmethoden und auch die Konfigurationsmöglichkeiten mit *iptables* unter Linux erläutert.

Bei der Konzeption der Übung für das Firewall-Modul erwies es sich als schwierig, offensive Übungen, z.B. zum Umgehen der Firewall, zu erstellen. Daher ist die Übung für beide Kurse identisch, in welcher die Teilnehmer ein vorgegebenes Szenario mit *iptables* implementieren sollen. So sind u.a. keine Verbindungen von außen, ausgenommen von SSH, erlaubt und zudem müssen alle verworfenen Pakete protokolliert werden. Zusätzlich sollen die Teilnehmer, die früh mit der Konfigurationsaufgabe fertig sind, eine kompliziertere Regel implementieren, die den Schutz vor *SYN-Flooding-Angriffen* gewährleistet.

4.2.7 Modul 7: Web-Anwendungssicherheit

Der Zugriff auf das Internet mit einem Webbrowser ist für Benutzer besonders leicht und komfortabel, doch auch hier bestehen verschiedene Sicherheitsrisiken. Den Teilnehmern werden daher nach einer allgemeinen Einführung in Webanwendungen verschiedene Angriffsmethoden mit den entsprechenden Gegenmaßnahmen vorgestellt, z.B. das

Einschleusen von Informationen mit *SQL Injections* in einen Datenstrom, das Umgehen von Zugriffsbeschränkungen durch *Broken Access Controls* oder das Auslesen von unzureichend geschützten Passwörtern.

Analog zum letzten Modul Firewalls konnte auch in der Übung zum Modul Webanwendungssicherheit keine Abgrenzung zwischen den beiden Kursen gefunden werden, da es nicht gelang, eine sinnvolle Übung für den Defensivkurs zu konstruieren. Die gemeinsame Übung besteht in der Anwendung von *WebGoat* [OWASP], einer kostenlosen Webanwendung, mit der durch praktische und interaktive Beispiele die oben beschriebenen Risiken und Sicherheitslücken von webbasierten Anwendungen sehr anschaulich vermittelt und geübt werden konnten.

4.2.8 Modul 8: Malware

Malware bezeichnet Computerprogramme, die unerwünschte und schädliche Funktionen ausführen. Der letzte Themenblock vor dem Abschlusstest behandelt die Klassifikation und wesentlichen Unterschiede von Viren, Würmern und Trojanischen Pferden sowie eine detaillierte Betrachtung der Funktionsweise von Rootkits mit den entsprechenden Abwehrmaßnahmen.

Der gemeinsame Teil der Übung besteht für die Versuchsteilnehmer darin, das Linux-Rootkit *LRK4* auf ihrem System zu installieren und die verschiedenen Funktionen des Rootkits, z.B. dem Verstecken von Dateien oder Prozessen, kennen zu lernen.

4.2.9 Modul 9: Abschlusstest

Bei dem letzten Themenblock handelt es sich um den Abschlusstest, der das wichtigste Messinstrument der Studie darstellt. Hierfür wird ein System mit zwölf Sicherheitslücken bzw. Konfigurationsfehlern präpariert; außerdem sind drei Aufgaben von den Teilnehmern zu bearbeiten. Eine Übersicht über alle Aufgaben und die erwarteten Ergebnisse ist im folgenden angegeben.

Malware

Offensiv	Defensiv
Detailliertere Behandlung der Rootkitfunktionen und Installation des Kernel-Rootkits Adore-NG	Initialisieren und Überprüfen der Tripwire-Datenbank.
Installation und Analyse eines Trojanischen Pferdes für versteckten Remote-Login	Suchen nach Rootkits Detaillierte Analyse des Quellcodes eines Wurms

Tabelle 4.8: Unterschiedliche Übungsinhalte des Moduls Malware

Teil 1 – Sicherheitslücken und Konfigurationsfehler:

1. **Rootkit (LRK)** Ein echtes Rootkit (Linux Rootkit, LRK) ist installiert und versteckt. Aufgabe der Teilnehmer ist es, das Vorhandensein des Rootkits mit einem Rootkit-Detektionsprogrammen wie *chkrootkit* [Chkrootkit] zu finden und korrekt als LRK zu identifizieren.

2. **Rootkit (Beastkit)** Es wird ein fingiertes Rootkit eingerichtet, indem eine Datei mit Namen „Arobia“ auf dem System angelegt wird. Das Vorhandensein dieser Datei ist charakteristisch für das Beastkit-Rootkit und wird vom Programm *Rootkithunter* [RKHunter] als solches identifiziert. Aufgabe für die Teilnehmer ist, anhand der Protokolldateien des *Rootkithunters* die Position der verdächtigen Datei ausfindig zu machen und anhand des Inhalts festzustellen, dass es sich nicht um ein reales Rootkit handelt. Sinn dieser Aufgabe ist festzustellen, ob die Teilnehmer Programmausgaben blind vertrauen.

3. **Root-User** Es existiert ein zusätzlicher Benutzer, welcher in der Datei *passwd* mit der User-ID 0, also mit Rootrechten, eingetragen ist. Von den Teilnehmern wird erwartet, diesen Fehler zu finden und den Benutzer zu löschen oder seine User-ID zu ändern.

4. **Sudo** In der Datei */etc/sudoers* ist ein nicht-berechtigter Benutzer eingetragen, der

mittels *sudo* die Befehle *kill* und *chmod* ausführen darf. Diese Rechte sollen dem Benutzer wieder entzogen werden.

5. **SUID Root** Die Texteditoren *nano* und *vim* sind SUID Root gesetzt. Aufgabe der Teilnehmer ist es, mit dem entsprechenden Kommando nach SUID-Root-Dateien zu suchen und anhand der Ausgabe die beiden genannten Programme als Sicherheitsrisiko zu identifizieren.

6. **Zugriffsrechte** Das */var/log*-Verzeichnis, das die Protokolldateien enthält, ist mit maximalen Zugriffsrechten versehen, wodurch jeder Benutzer Protokolleinträge verändern kann. Dieses Problem soll erkannt und die korrekten Zugriffsrechte (kein Schreiben für unprivilegierte Benutzer) gesetzt werden.

7. **Schwache Passwörter** Mehrere Benutzer des Systems haben sehr schwache Passwörter. Die Teilnehmer sollen diese Passwörter mittels eines Passwort-Crackers ermitteln und die Einhaltung einer auf dem Aufgabenblatt vorgegebenen Passwortrichtlinie überprüfen.

8. **Serverdienste** Aus dem Aufgabenblatt geht hervor, dass als Netzwerkdienste nur SSH und Telnet auf dem System laufen dürfen, es sind aber zusätzlich noch ein FTP-Server und der Finger-Daemon aktiv. Von den Teilnehmer wird erwartet, die offenen Ports und Dienste zu identifizieren und so zu deaktivieren, dass sie auch bei einem Neustart des Systems nicht mehr gestartet werden. Eine temporäre Lösung z.B. durch Verwendung des *kill*-Befehls gibt daher nicht die volle Punktzahl.

9. **SYN-Cookies** In der Datei */etc/network/options* ist die Verwendung von SYN-Cookies auf „no“ gesetzt, wodurch das System anfällig ist gegen SYN-Flood--Angriffe. Für eine korrekte Bearbeitung der Aufgabe muss dieser Wert auf „yes“ gesetzt werden.

10. **Systemintegrität** Die *Tripwire*-Datenbank enthält Hinweise auf weitere Kompromittierungsversuche, z.B. der versuchten Einrichtung des Rootkits *Sebek*.

11. Tcpdump Die Teilnehmer sollen feststellen, dass auf dem System ein Netzwerk-Sniffer aktiv ist und diesen beenden.

12. Promiscuous Mode Es solle erkannt werden, dass eine Netzwerk-Schnittstelle des Systems in den promiscuous mode versetzt ist, sodass alle an dieser Schnittstelle ankommenden Daten verarbeitet werden können, und nicht nur die an das Gerät selbst gerichteten Pakete. Dies deutet auf einen laufenden Netzwerk-Sniffer hin.

Bei den ausgewählten Problemen handelt sich um

- die direkte Anwendungen von Erlerntem aus dem praktischen Teil des Kurses (1., 3., 4., 5., 7., 9., 10.),
- Sachverhalte, die nur im theoretischen Teil erwähnt wurden (6., 8.) und
- Transferaufgaben, die die Übertragung von Wissen und neues Wissen erforderten (2., 9., 11., 12.).

Teil 2 – Aufgaben:

Aufgabe 1: Firewall-Einstellungen Die Teilnehmer sollen die Übereinstimmung der Konfiguration einer Paketfilter-Firewall, die als Personal-Firewall arbeitet, mit den Vorgaben auf dem Aufgabenblatt überprüfen und ggfs. korrigieren. In der Firewall-Konfiguration sind zum einen die Firewallrichtlinien (Accept, Drop) falsch gesetzt, zum anderen ist der DNS-Port mit „63“ angegeben, anstatt korrekt mit „53“.

Aufgabe 2: SSH Die zweite Aufgabe der Teilnehmer besteht darin, SSH nach den Angaben auf dem Aufgabenblatt zu konfigurieren. So soll u.a. kein direktes Einloggen als Benutzer Root möglich sein und jeder Benutzer soll zur Eingabe eines Passwortes gezwungen werden.

Aufgabe 3: Sudo Einem Benutzer soll die Verwendung des *Mount*-Befehls per *sudo* erlaubt werden.

4.2.10 Anmerkungen

Eine eindeutige Abgrenzung zwischen den beiden Lehransätzen zu finden erwies sich teilweise als sehr schwierig, da viele Administratoren selbst „Hacker-Tools“ wie *nmap*, *Nessus* oder *John the Ripper* einsetzen, um die von ihnen betreuten Systeme zu überwachen. Die Übungen enthielten daher auch immer einen – teilweise auch recht großen – gemeinsamen Aufgabenteil.

4.3 Durchführung der Studie

Die Studie wurde im Rahmen dieser Arbeit zweimal durchgeführt: einmal im Frühjahr 2007 an der RWTH Aachen und einmal im Frühjahr 2008 an der Universität Mannheim und der RWTH Aachen. In dieser Arbeit wird nur die erste Durchführung vorgestellt, für Informationen zur zweiten siehe die Angaben in Abschnitt 6.1.

Die Durchführung der beiden Kompaktkurse im Jahr 2007 erfolgte vom 20.–22. März sowie vom 27.–29. März im Rechenzentrum der RWTH Aachen. Der dortige PC-Pool bot aufgrund seiner Größe und Ausstattung gute Voraussetzungen für die Veranstaltung.

4.3.1 Auswahl der Stichprobe und der Gruppen

Die Kurse wurden auf der Homepage des Lehrstuhls und über diverse Mailinglisten angekündigt. Die Anmeldung erfolgte per E-Mail durch Einsendung des ausgefüllten Fragebogens (d.h. dem Awareness- und dem Wissenstest) und persönlicher Angaben (u.a. Name, Alter, Universität, Fachsemester und einem Motivationstext), siehe Anhang A. Innerhalb kurzer Zeit gingen über 100 Bewerbungen von Studenten aus ganz Deutschland ein, darunter auch Anfragen von Firmen für ihre Mitarbeiter. Aufgrund räumlicher und untersuchungstechnischer Bedingungen standen für das Experiment jedoch maximal 48 Plätze zur Verfügung, wodurch mehr als die Hälfte der Bewerbungen nicht berücksichtigt werden konnte. Die Auswahl der Teilnehmer erfolgte nicht zufällig, sondern nach zuvor festgelegten Kriterien:

- Die Kompaktkurse sollten bereits im Vorjahr angeboten werden, mussten aus zeit-

lichen Gründen aber abgesagt werden. Allen Bewerbern, die bereits für das Vorjahr eine Zusage erhalten hatten, wurde ein Vorrecht auf einen Platz im Kurs eingeräumt.

- Da die Stichprobe ein hohes Maß an Repräsentativität für die Population der Studierenden aufweisen sollte, wurden nur Studenten mit einem Alter von maximal 30 Jahren zugelassen.
- Des Weiteren richtete sich die Auswahl der restlichen Teilnehmer danach, dass in jeder Vorwissensgruppe etwa gleich viele Teilnehmer waren, um die Ergebnisse nicht durch ungleich große Stichproben zu verzerren.
- Konnte zwischen der Auswahl zweier Teilnehmer nach den obigen Kriterien nicht entschieden werden, erhielt der Teilnehmer eine Zusage, der sich zuerst für die Kurse angemeldet hatte.
- Teilnehmer, die nicht berücksichtigt werden konnten, rückten automatisch in der Reihe ihrer Anmeldung auf die Warteliste, um im Falle einer Absage kurzfristig nachrücken zu können.

Tabelle 4.9 zeigt eine Übersicht über die ausgewählten Teilnehmer bezüglich der vertretenen Hochschulen und Fachrichtungen sowie dem Durchschnittsalter und Durchschnittssemester aller Teilnehmer. Wie zu erkennen ist, waren nicht nur Informatik-Studiengänge vertreten, was darauf hinweist, dass das Interesse an IT-Sicherheit allgemein hoch ist.

4.3.2 Zuordnung der Teilnehmer zu den Gruppen

Die Zuordnung in die Vorwissensgruppe erfolgte anhand des Wissenstests des ersten Fragebogens. Als Grenze wurden 20 Punkte, d.h. knapp die Hälfte der maximal erreichbaren Punkte gewählt. Probanden, die im Wissenstest 20 Punkte oder mehr erreicht haben, wurden in die Gruppe mit viel Vorwissen eingeteilt, Teilnehmer mit weniger Punkten in die Gruppe mit wenig Vorwissen. Die nachfolgende Aufteilung auf die beiden Kurse

Universitäten (Anzahl)	Studiengänge (Anzahl)	Alters-Durchschnitt	Semester-Durchschnitt
Aachen (36) Konstanz (1) Oldenburg (1) Bremen (1) Berlin (1) Bochum (1) Köln (1)	Informatik Diplom (27) Informatik B.Sc (1) Informatik Lehramt (2) Informationstechnik (1) Mathematik (1) BWL (Nebenfach Informatik) (1) Maschinenbau (1) Technische Kommunikation (1) Medienwissenschaften (1) Elektrotechnik (1)	24,1	7,33

Tabelle 4.9: Übersicht über die Teilnehmer

fand zufällig statt. Abbildung 4.2 stellt den Vorgang anschaulich dar. Hierbei bezeichnet „EG“ die Experimentalgruppe und „KG“ die Kontrollgruppe.

4.3.3 Organisation und Technik

Ein Kurstag dauerte von 9.30 bis 17.00 Uhr. Pro Tag fanden drei Module statt, mit einer Mittagspause von 11.30 bis 12.45 Uhr und einer weiteren Pause von 15.45 bis 15.00 Uhr. Die Teilnehmer-PCs waren untereinander mit Hubs verbunden, um im Netzwerkmodul ein einfaches Netzwerksniffen zu erlauben. Über einen Switch bestand darüber hinaus eine Anbindung ans Internet, was für einige Übungen erforderlich war. Während der Arbeit mit potentiell problematischen Anwendungen wie DoS[2]-Tools wurde die Verbindung vom Switch zum Rechenzentrumsnetz getrennt, sodass das Netzwerk der Kursteilnehmer von anderen Netz isoliert war. Zusätzlich existierte für verschiedene Aufgaben ein Server, der so genannte Kursserver.

Damit die Teilnehmer nicht direkt auf der Hard- und Software der Kursrechner arbeiten, sie mit Root-Rechten arbeiten können sowie die Installation einfacher vorbereitet und aufgespielt werden kann, wurde eine Virtualisierungssoftware verwendet. Es

[2]Denial of Service

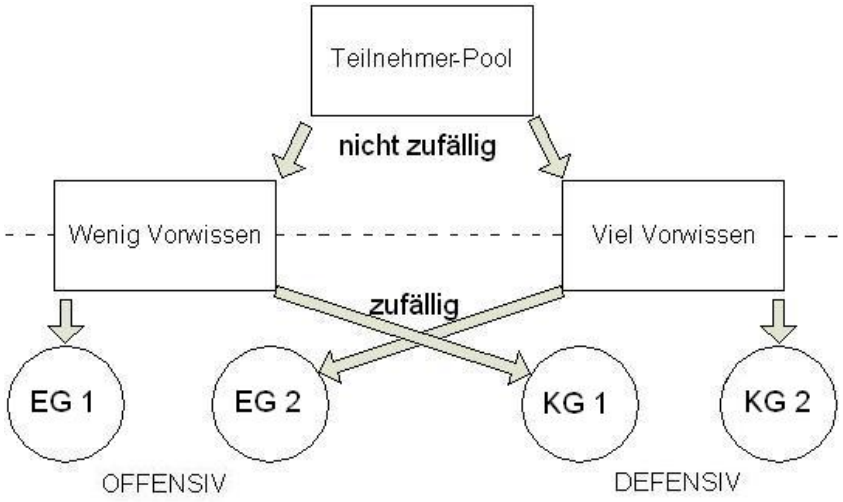


Abbildung 4.2: Auswahl und Zuordnung der Teilnehmer

wurde *Virtual PC 2007* von Microsoft eingesetzt, welches die einzige in dem PC-Pool nutzbare Virtualisierungssoftware und auf allen Rechnern bereits vorinstalliert war. Als Betriebssystem wurde von den Versuchsleitern ein *Linux Debian System* gewählt und entsprechend für die Kurse vorkonfiguriert, welches als virtuelle Festplatte auf alle Kursrechner verteilt wurde und jeder Versuchsteilnehmer mit einem identischen System arbeiten konnte. Linux wird zwar offiziell nicht von Microsoft Virtual PC 2007 unterstützt, abgesehen von ein paar Einschränkungen ist die Nutzung jedoch möglich.

Nach dem Vortrag jedes Moduls konnten sich die Teilnehmer die Vortragsfolien und des Übungsblatt von einem Webserver des Kursservers herunterladen. Am Ende eines Moduls sollten sie – in Ermangelung einer bequemeren Alternative – das bearbeitete Aufgabenblatt sowie die bearbeiteten Dateien mittels SCP[3] in ein ihnen zugeordnetes Verzeichnis hochladen. Auf dem Aufgabenblatt konnten die Teilnehmer außerdem Feedback zu der jeweiligen Übung angeben.

Im folgenden werden der Awarenesstest und der Wissenstest – beides Bestandteile des Fragebogens – sowie der Abschlusstest vorgestellt.

[3]Secure CoPy; Kommandozeilenprogramm zum Übertragen von Dateien über eine verschlüsselte Verbindung

4.3.4 Awarenesstest

Für die Konstruktion der Testitems wurde ein gebundenes Aufgabenformat gewählt, es wurden also verschiedene Antwortalternativen für jede Frage vorgegeben. Neben einfachen Ja/Nein-Behauptungen wie „Ich verschlüssele meine E-Mails“ wurden in dem Fragebogen auch Ratingformate verwendet. Daher wurde, um ein unsicheres Verhalten stärker zu gewichten, für viele Fragen eine bipolare Ratingskala verwendet, für welche sowohl negative als auch positive Werte möglich waren. Im Awarenesstest konnten somit von -12 bis $+24$ Punkten erzielt werden, wobei ein hoher Punktwert eine hohe Sensibilisierung des Teilnehmers gegenüber Informationssicherheit repräsentierte. Der vollständige Fragebogen ist in Anhang A.2 zu finden.

Nach dem Kurs erhielten die Teilnehmer einen ähnlichen Test, in welchem viele Fragen aus dem Vortest erneut aufgegriffen wurden. Dabei wurden die Items nicht mehr als Behauptungen formuliert, sondern als konkrete Fragen, um die Teilnehmer nach der Bereitschaft zur Veränderung ihres bestehenden Verhaltens zu befragen, z.B. ob sie nach den Erfahrungen aus dem Kurs ihre E-Mails oder Daten verschlüsseln oder häufiger Backups durchführen würden. Die Fragen dieses Tests sind in Anhang B zu finden.

4.3.5 Wissenstest

Dieser Wissenstest war als Niveautest aufgebaut, der Test begann also mit sehr leichten Fragen und wurde zum Ende hin schwerer. Dabei wurden leichte Fragen bei korrekter Antwort mit einem Punkt gewertet, mittelschwere Fragen mit zwei Punkten und schwere Fragen mit drei Punkten, wodurch bei 19 Fragen insgesamt 41 Punkte im Wissenstest erreicht werden konnten. Anhand der Ergebnisse dieses Tests wurden die Teilnehmer bezüglich ihrer Vorkenntnisse eingestuft.

Um die Wahrscheinlichkeit, Antworten bei Unwissenheit richtig zu raten, zu minimieren, wurde der Test als Multiple-Choice-Test aufgebaut, das heißt, es wurden pro Frage drei bis sechs Antwortalternativen vorgegeben. Um die Wahrscheinlichkeit korrekter Antworten durch Raten zu verringern, war nicht angegeben, wie viele der Antwortalternativen korrekt sind. Die Bewertungsrichtlinien sahen für diesen Test vor,

- für jede falsch markierte Antwort einen Punkt abzuziehen,
- für nicht markierte, korrekte Antworten oder die Antwort „Weiß ich nicht“ weder einen Punkt zu vergeben noch einen abzuziehen und
- jede Aufgabe mit mindestens 0 Punkten zu bewerten.

Da der Wissenstest als Niveautest aufgebaut war, sollte die Schwierigkeit der Fragen im Verlauf des Tests zunehmen. Wie schwer eine Frage für die Teilnehmer zu beantworten war, kann letztendlich durch die Berechnung eines Schwierigkeitsindex für jedes Testitem beobachtet werden. Unter Verwendung von Formel 3.1 in Abschnitt 4.1.4 ergibt sich für den Schwierigkeitsindex des Wissenstests der in Abbildung 4.3 dargestellte Graph. Wie dem Graph zu entnehmen ist, nahm insgesamt die Wahrscheinlichkeit einer korrekten Beantwortung der Testitems mit steigendem Fragenniveau ab. So lässt sich z.B. feststellen, dass Frage 11 „Wofür steht die Abkürzung CIA in der IT-Sicherheit?“ und Frage 16 „Welches Programm kann für einen Distributed Denial of Service-Angriff verwendet werden?“ für die Teilnehmer am Schwierigsten zu beantworten waren. Der vollständige Wissenstest ist in Anhang A.3 zu finden.

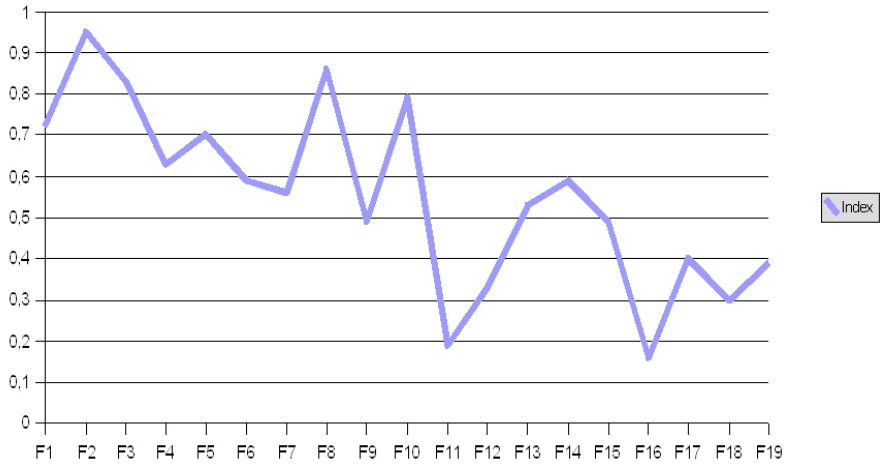


Abbildung 4.3: Bestimmung der Itemschwierigkeit des Wissenstest

Am Ende des Kompaktkurses erhielten die Versuchsteilnehmer denselben Test erneut zur Beantwortung. Da die Lösungen des Vortests nicht besprochen wurden, konnte hier

ein direkter Vergleich zwischen den Ergebnissen vor und nach der Kursdurchführung erfolgen.

4.3.6 Entwicklung und Evaluation des Awareness- und des Wissenstests

Um missverständliche oder zu leichte bzw. zu schwere Fragen bereits im Vorfeld der Untersuchung auszuschließen, wurde der Wissens- und Persönlichkeitstest einer nicht am Experiment beteiligten Stichprobe, hier einer Gruppe von 30 Fachhochschulstudenten, vor der Kursdurchführung als Probedurchlauf vorgelegt.

Für den Wissenstest ergibt sich mit Formel 4.1 (siehe Abschnitt 4.1.4) ein Alphakoeffizient von $\alpha = 0,697$. Der Wissenstest hat also eine Messgenauigkeit von knapp 70%. Zur Erinnerung: ein Messwert setzt sich in der klassischen Testtheorie aus dem wahren Wert einer Ausprägung und einem Messfehler zusammen (siehe Abschnitt 3.1.5). Daraus folgt, dass der Messfehler hier einem Wert von ungefähr 30% entspricht. Über die Genauigkeit dieses Ergebnisses kann aber nur spekuliert werden, da zum einen die Anzahl der Fragen relativ klein war und zum anderen auch die Stichprobe nicht ausreichend groß bzw. mit etwa 40 Teilnehmern weit unter dem von Mendoza u. a. [2000] geforderten Minimum von 100 Testpersonen lag.

4.3.7 Abschlusstest

Für den Abschlusstest wurde das im Kurs verwendete System als Basis genommen und mit den in Abschnitt 4.2.9 vorgestellten 12 Sicherheitslücken bzw. Fehlkonfigurationen präpariert. Die Verwendung des Systems aus den ersten acht Modulen hat für die Teilnehmer den Vorteil, dass das System, die Bedienung und die vorhandenen Programme bereits bekannt sind; für den Veranstalter hat es den Vorteil, dass weniger Arbeits- und Vorbereitungsaufwand nötig ist.

Die Teilnehmer erhielten vor Beginn des Tests ein Aufgabenblatt mit Instruktionen (siehe Anhang C), auf dem die Aufgabenstellung beschrieben war. Als Situation war gegeben, dass der Teilnehmer die Administration eines Computers in einer Firma

übernimmt, von dem er nur wenige Informationen hat. Informationen waren angegeben zu den angebotenen Serverdiensten, der Passwortrichtlinie sowie der (beabsichtigten) Konfiguration der Firewall. Die Aufgabenstellung lautete, den Computer auf Sicherheitslücken und Fehlkonfigurationen zu untersuchen und in einen sicheren Zustand zu bringen; außerdem wurden drei Aufgaben gestellt, die nach abgeschlossener Analyse des Systems bearbeitet werden sollten.

Mit Beginn des Tests erhielten die Teilnehmer die Zugangsdaten zum Einloggen, und die Startzeit wurde notiert. Das Kursmaterial (Folien und Übungsblätter) durfte während des Tests verwendet werden. Am Ende des Tests wurden die Teilnehmer aufgefordert, das Aufgabenblatt sowie die veränderten (Konfigurations-)Dateien auf den Kursserver hochzuladen.

Protokollierung

Für die spätere Auswertung des Tests war es nötig, das Vorgehen sowie die Veränderungen durch die Teilnehmer nachvollziehen zu können. Zu diesem Zweck wurde das Vorgehen der Probanden protokolliert. Da der Keylogger (siehe weiter unten) nur eingeschränkte Ergebnisse erwarten ließ, waren die Teilnehmer aufgefordert, ihr Vorgehen auf dem Aufgabenblatt des Abschlusstests zu protokollieren, indem sie die identifizierten Probleme, die Lösungsvorschläge zu deren Behebung sowie die dafür verbrauchte Zeit eintrugen.

Zur Protokollierung des Vorgehens wurden folgende Möglichkeiten identifiziert: Keylogger, Einträge in den Log-Dateien, Protokolldateien (u.a. die verwendeten Kommandozeilenbefehle in der Datei `bash_history`). Ein Keylogger schreibt alle Tastatureingaben des Benutzers mit und bietet so eine elegante Möglichkeit, das Vorgehen des Benutzers zu protokollieren. Allerdings erwies sich dies in der Praxis als nicht einfach, da die recherchierten, frei verfügbaren Keylogger entweder alle Eingaben oder nur die in einem Terminalprogramm aufzeichnen. Im ersten Fall besteht das Problem, dass auf einem System mit grafischer Oberfläche später nicht nachvollziehbar ist, in welchem Fenster bzw. welcher Anwendung die Eingaben anfielen; im zweiten Fall, dass nur die Eingaben der Kommandozeile aufgezeichnet werden. Außerdem ist der Nachteil eines Keyloggers, dass auf einem System mit grafischer Oberfläche Eingaben durch Mausklicks nicht protokol-

liert werden. Da für den Abschlusstest – bedingt durch zu gering vorhandenen Speicherplatz (siehe Abschnitt 5.5.2) – ein System ohne grafische Oberfläche verwendet werden musste, wurde ein Keylogger für das Terminalprogramm gewählt, der in vielen Linux-Distributionen standardmäßig installiert ist: `script`. Damit jedoch später nachvollziehbar ist, welchem Terminal bestimmte Eingaben zuzuordnen sind, darf die Protokollierung aller geöffneten Terminalsitzungen nicht in eine Datei (die Standarddatei) erfolgen, sondern für jedes Terminal getrennt. Dafür muss `script` mit der entsprechenden Option gestartet werden (ein entsprechender Hinweis befand sich auf dem Aufgabenblatt). Ein Nachteil von `script` ist, dass die Daten erst beim regulären Beenden der jeweiligen Shell in die Protokolldatei geschrieben werden; wird die Shell abgebrochen, gehen die Änderungen verloren. Zur Unterstützung der späteren Auswertung wurden Systembestandteile, in denen potentiell Spuren hinterlassen wurden, wie die bereits genannte Datei `bash_history`, die Konfigurationsdateien im Verzeichnis `/etc/`, die Log-Dateien im Verzeichnis `/var/log/` sowie die Protokolldateien von `script`, gesichert.

4.4 Erfahrungen

Bereits während der Konzeption der Kurse hatte sich gezeigt, dass eine Abgrenzung von offensiven und defensiven Methoden nicht einfach und nicht eindeutig möglich ist. Zusätzlich bestand das Problem, dass beide Gruppen trotz der Differenzierung am Ende das nötige Wissen haben müssen, um die Aufgaben des Abschlusstests bearbeiten zu können. Als schwierig erwies sich auch die Protokollierung des Abschlusstests für die Testauswertung. Es ist nicht trivial, ein System zu finden, das den Anforderungen genügt, das Vorgehen der Teilnehmer sowie die von ihnen verursachten Veränderungen aufzuzeichnen und zusätzlich eine möglichst wenig zeitaufwändige (d.h. automatisierte) Auswertung ermöglicht. Außerdem stellte sich später bei der Auswertung heraus, dass einige Teilnehmer den Inhalt der Datei `bash_history` gelöscht hatten, mit dem sich die Abfolge der verwendeten Kommandozeilenbefehle nachvollziehen lässt.

Während der Durchführung der Kurse zeigten sich große Unterschiede in der Bearbeitungszeit des Praxisteils zwischen Teilnehmern mit und solchen ohne Vorkennt-

nissen – speziell in Systemadministration, aber auch IT-Sicherheit, Programmierung und Netzwerke. Unterschiedliche lange Bearbeitungszeiten war bereits beim Entwurf der Übungsblätter berücksichtigt worden, indem Zusatzaufgaben angeboten wurden. Trotzdem passierte es jedoch, dass einzelne Teilnehmer die Zusatzaufgaben bearbeitet hatten, während andere noch nicht mit den von allen zu bearbeitenden Aufgaben fertig waren. Hier zeigten sich auch Probleme in Bezug auf die Gegebenheiten in den PC-Pools und die Verwendung von Virtualisierungssoftware, auf die in Abschnitt 5.5.2 detaillierter eingegangen wird.

Ein weiteres Problem war die Motivation der Teilnehmer. An der Besprechung der Übungsaufgaben beteiligten sich nur wenige aktiv und meist erst nach mehrmaliger Aufforderung. Es gab Teilnehmerschwund dadurch, dass Teilnehmer erst gar nicht den Kurs begannen oder während der Kurslaufzeit absprangen. Und einige Teilnehmer nahmen nicht durchgängig an dem Kurs teil.

Positiv war die Unterstützung durch die Betreiber der für die Kurse verwendeten PC-Pools an der RWTH Aachen und der Universität Mannheim.

4.5 Zusammenfassung

In diesem Kapitel wurde die Konzeption der Studie, der Entwurf der Kompaktkurse zur Verwendung in der Studie sowie die Durchführung der Studie vorgestellt. Ein wesentlicher Bestandteil der Konzeption der Studie war die Konstruktion geeigneter Tests zur Messung des IT-Sicherheitsverständnis, welche durch standardisierte Bewertungsmaßstäbe eine objektive Messung der Ergebnisse garantieren sollen, um die beiden Lehransätze miteinander vergleichen zu können. Der schwierigste Teil der Planung bestand in der Konzeption der Übungen zu den ausgewählten Themengebieten, da eine Abgrenzung zwischen offensiven bzw. defensiven Inhalten oft nicht eindeutig war und letztendlich auch nicht bei jeder Übung gelungen ist.

5 Ergebnisse

Nachdem im vorherigen Kapitel die Vorbereitung der Studie und die Durchführung der beiden Kompaktkurse beschrieben wurde, erfolgt in diesem Kapitel die Auswertung aller erhobenen Daten und die Interpretation der Ergebnisse. Dies beinhaltet auch die Durchführung eines *Signifikanztests*, um zu überprüfen, welche Aussagekraft den Ergebnissen beigemessen werden kann und wie diese zu interpretieren sind. Das Kapitel endet mit einer kritischen Reflexion der Studie und einem Ausblick. Die Auswertung der Studie erfolgte überwiegend mit der Statistiksoftware SPSS [SPSS]

5.1 Auswertung der Tests

In den folgenden Abschnitten wird das Vorgehen bei der Auswertung vorgestellt und wie die Punktevergabe erfolgte.

5.1.1 Awarenesstest und Wissenstest

Die Papierfragebögen wurden in SPSS erfasst und dann mit den von SPSS angebotenen Methoden ausgewertet.

Die Punktevergabe erfolgte wie in den Abschnitten 4.3.4 und 4.3.5 vorgestellt: im Awarenesstest von -12 bis $+24$, im Wissenstest von 0 bis 41 Punkte. Ein hoher Punktwert im Awarenesstest repräsentiert eine hohe Sensibilisierung des Teilnehmers bezüglich Informationssicherheit.

5.1.2 Abschlusstest

Durch die Auswertung der Aufzeichnungen der Teilnehmer, der von den Teilnehmern veränderten Dateien, der Protokolldateien des Keyloggers, der Datei `bash_history` und der System-Logs wurde nachvollzogen, welche Aufgaben die Probanden bearbeitet und wie sie das jeweilige Problem gelöst hatten.

Für jede Aufgabe des Tests (siehe Abschnitt 4.2.9; hier sind auch die erwarteten Lösungen angegeben) wurde maximal ein Punkt vergeben, es konnten im Abschlusstest also insgesamt 15 Punkte erzielt werden. Teilweise gelöste Aufgaben bzw. die Angabe von unzureichenden Lösungsvorschlägen für gefundene Probleme wurden mit einem halben Punkt gewertet.

5.2 Ergebnisse des Abschlusstests

Der primäre Schritt zur Überprüfung der Hypothese und damit zum Vergleich der beiden Lehransätze ist die Auswertung des praktischen Abschlusstests. Wie bereits in der Planung des Abschlusstests (siehe Abschnitt 4.2.9) beschrieben, sollte in diesem Test die Anzahl der gefundenen Sicherheitslücken und Konfigurationsfehler gemessen werden.

5.2.1 Auswertung der gefundenen Sicherheitslücken

Die Ergebnisse des Abschlusstests ohne eine Unterscheidung nach Vorwissen sind in Abbildung 5.1 dargestellt. Die y-Achse entspricht den erreichten Punkte aus dem Test, deren genaue Werte Tabelle 5.1 entnommen werden können.

	Punkte	in %	Konfidenzintervall
Defensivkurs	5,7	38,07%	[4,7 ; 6,75]
Offensivkurs	6,11	40,73%	[5,2 ; 7]

Tabelle 5.1: Ergebnisse des Abschlusstest

Wie an diesen Ergebnissen zu erkennen ist, haben die Teilnehmer des Offensivkurses im Abschlusstest mit durchschnittlich 40,73% der maximal erreichbaren Punkte ein

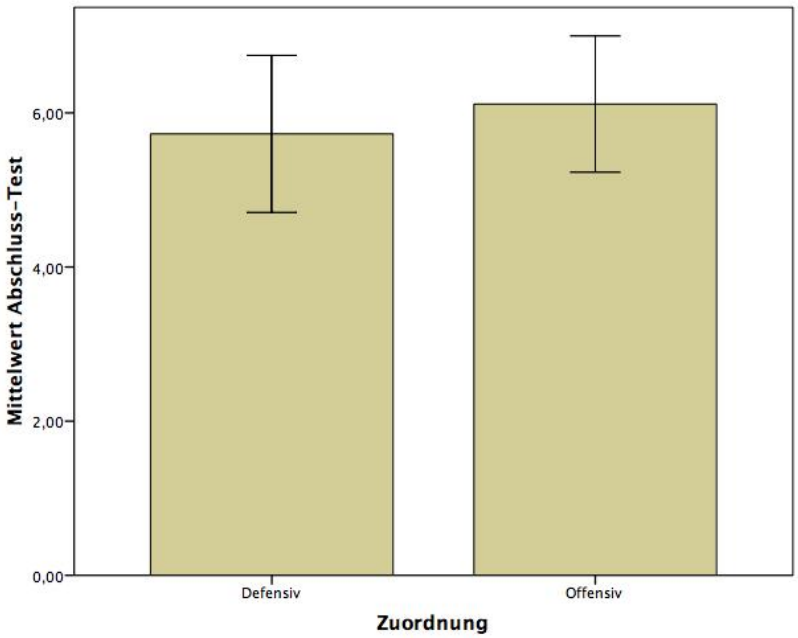


Abbildung 5.1: Ergebnisse des Abschlusstest

besseres Ergebnis erzielt als die Teilnehmer des Defensivkurses mit 38,07%. Anhand dieser Resultate allerdings zu folgern, dass der offensive Lehransatz besser sei als der defensive Ansatz, wäre zum einen verfrüht und zum anderen mit einer hohen Irrtumswahrscheinlichkeit verbunden, wie es sich im Verlaufe dieses Kapitels zeigen wird. Es ist offensichtlich, dass der Unterschied, der gerade einmal einer Differenz von 0,41 Punkten im Abschlusstest entspricht, äußerst knapp ist. Hinzu kommt die Unschärfe des Ergebnisses, welche durch eine Betrachtung der Konfidenzintervalle festgestellt werden kann. Ein *Konfidenzintervall* – auch als Vertrauensintervall bezeichnet – schätzt den Bereich, in welchem der Mittelwert der Testresultate mit einer bestimmten Wahrscheinlichkeit (häufig 95%) liegt. Je kleiner dieses Intervall ist, desto präziser ist der Mittelwert und somit auch das Testergebnis. Die Konfidenzintervalle für 95% sind im Diagramm in Abbildung 5.1 eingezeichnet (Linien mit beiderseitigem Abschluss). Wie jedoch zu erkennen ist, sind die Intervalle relativ breit und überlappen sich zu einem großen Teil. Da die Lage des Mittelwertes nicht genauer bestimmt werden kann, handelt es sich um ein nicht

signifikantes, also nicht aussagekräftiges Testergebnis.

Unter Berücksichtigung des Vorwissens der Teilnehmer (siehe Abbildung 5.2) wird ersichtlich, dass die Teilnehmer in den Gruppen mit wenig Vorwissen durchschnittlich fast identische Ergebnisse erzielt haben, wobei der Wert des Defensivkurses geringfügig besser ist. In den Gruppen mit viel Vorwissen ist jedoch ein deutlicher Unterschied erkennbar, dass die Studenten des Offensivkurses mit durchschnittlich 6,82 Punkten rund 11% mehr Fehler im Abschlusstest gefunden haben als die Teilnehmer der Kontrollgruppe mit 5,96 Punkten (siehe Tabelle 5.2). Somit kann nicht ausgeschlossen werden, dass das als Kontrollvariable erhobene Vorwissen einen signifikanten Einfluss auf die Ergebnisse des Abschlusstests hat und wird daher bei den durchgeführten Signifikanztests in Abschnitt 5.4 genauer untersucht.

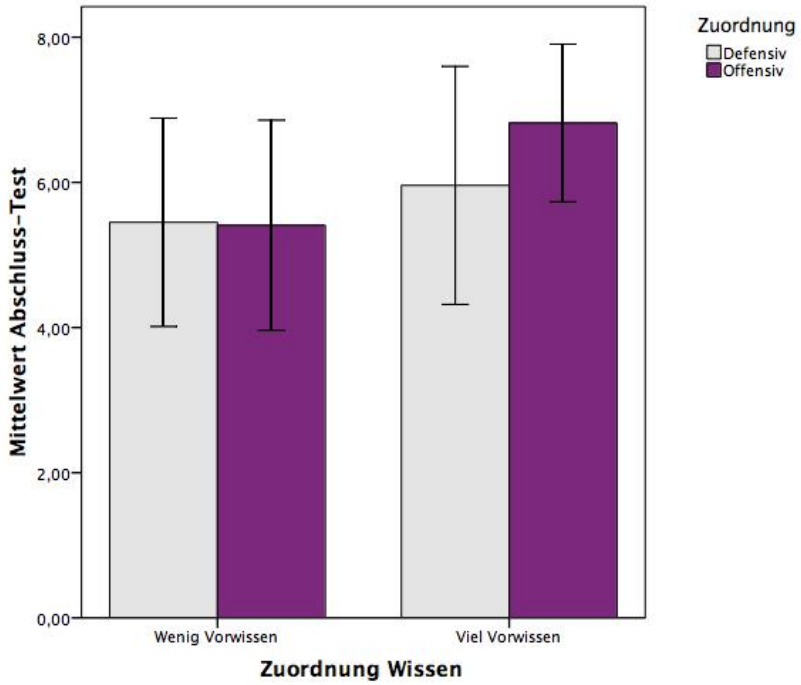


Abbildung 5.2: Ergebnisse des Abschlusstests nach Vorwissen

	Punkte	in %	Konfidenzintervall
Defensivkurs – Wenig Vorwissen	5,45	36,33%	[4 ; 6,8]
Defensivkurs – Viel Vorwissen	5,96	39,73%	[4,3 ; 7,6]
Offensivkurs – Wenig Vorwissen	5,41	36,07%	[3,9 ; 6,85]
Offensivkurs – Viel Vorwissen	6,82	45,47%	[5,7 ; 7,9]

Tabelle 5.2: Ergebnisse des Abschlusstest nach Vorwissen

5.2.2 Auswertung der Strategie

Die beiden Kurse sollten aber nicht nur nach der Anzahl identifizierter Probleme auf dem System verglichen werden, sondern auch anhand der Strategien, welche die Teilnehmer bei der Analyse des Systems verfolgt haben. Diese Vorgehensweise konnte sowohl anhand der Aufzeichnungen der Teilnehmer aus dem Abschlusstest als auch durch eine Analyse der Protokolldateien des Keyloggers rekonstruiert werden. Die „durchschnittliche“ Vorgehensweise aller Teilnehmer wurde nach dem folgenden Schema berechnet:

1. Jedem zu identifizierenden Problem wurden Punkte von 1 bis 15 entsprechend der Position, an deren Stelle dieses Problem von den Teilnehmern identifiziert wurde, zugewiesen.

   Beispiel: Die Überprüfung der eingetragenen Benutzer auf schlechte Passwörter wird von zwei Teilnehmern an erster Stelle durchgeführt, von einem dritten Teilnehmer an vierter Stelle. Dann erhielt das Problem „Schwache Passwörter“ den Wert $1 + 1 + 4 = 6$.

2. Die Ergebnisse aus dem ersten Schritt müssen verfeinert werden, da für jedes Problem auch die Anzahl Teilnehmer, die dieses Problem *nicht* entdeckt haben, in die Berechnung der durchschnittlichen Vorgehensweise mit einfließen. Daher erhielt jedes Problem für jeden Teilnehmer, der es nicht identifiziert hat, einen zusätzlichen Wert von 16, da die Werte von 1 bis 15 für die gefundenen Probleme entsprechend ihrer Position vorgesehen waren.

   Beispiel: Das Problem des aktiven Sniffers wird nur von einem einzigen Teilnehmer an erster Position gefunden, von den anderen 21 Teilnehmern des Kurses gar

nicht. Dadurch ergibt sich ein Wert für das Snifferproblem von $1 + 21 * 16 = 337$.

3. Die Reihenfolge der bearbeiteten Probleme ergab sich aus den Ergebnissen der so verfeinerten Werte. Je kleiner der berechnete Wert für ein Problem war, desto häufiger und früher wurde es im Durchschnitt von den Teilnehmern identifiziert. Bei identischen Werte wurde das Problem zuerst gewählt, das von mehr Teilnehmer gefunden wurden.

Für den Offensiv- und Defensivkurs wurden anhand des obigen Verfahrens die folgenden Strategien für die in Abschnitt 4.2.9 beschriebenen Probleme des Abschlusstests berechnet. In Klammern steht jeweils der nach dem Verfahren berechnete Wert.

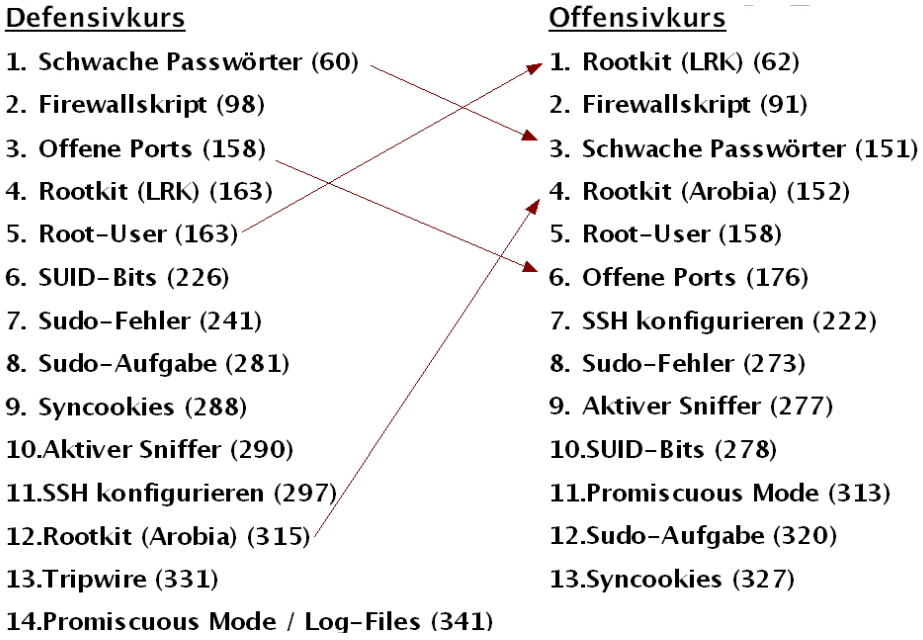


Abbildung 5.3: Bearbeitungsstrategien im Abschlusstest

An der Darstellung der unterschiedlichen Vorgehensweisen in Abbildung 5.3 ist zu erkennen, dass die Teilnehmer des Defensivkurses früher nach Benutzern mit schwachen Passwörtern gesucht und das Problem der offenen Ports gelöst haben. Dem gegenüber steht die Bearbeitungsstrategie der Offensivteilnehmer, welche häufiger zuerst nach Malware, also den beiden Rootkits, auf dem System suchten. Zudem identifizierten sie auch

häufiger und früher die Probleme des aktiven Sniffers und der Schnittstelle im promiscuous mode. Auffallend ist, dass die Versuchsteilnehmer des Defensivkurses – obwohl sie die Kenntnisse hatten – erst relativ spät nach Rootkits suchten. Dabei gehört die Verbreitung von Malware oder das Ausspähen von Informationen zu den größten realen Gefahren für IT-Systeme, was zu der Vermutung führt, dass die Teilnehmer des Offensivkurses eher die konkreten Bedrohungspotentiale für IT-Sicherheit einschätzen können und nach diesen suchen. Diese Vermutung wird auch bestätigt durch die Ergebnisse des Awarenesstests, welche im folgenden Abschnitt präsentiert werden.

5.3 Ergebnisse des Awareness- und des Wissenstests

Die Teilnehmer wurden anhand ihrer Resultate im Wissenstest vor Kursbeginn in die Subgruppen mit viel bzw. wenig Vorwissen eingeteilt. Wie Abbildung 5.4 und Tabelle 5.3 zeigen, ist die Ausgangslage von beiden Kursen einschließlich der Konfidenzintervalle aufgrund der Zuordnung nahezu gleich, was einen optimalen Vergleich mit den Ergebnissen nach der Kursdurchführung ermöglicht.

	Punkte	in %	Konfidenzintervall
Defensivkurs – Wenig Vorwissen	14	34,15%	[12,0 ; 16,0]
Defensivkurs – Viel Vorwissen	25,82	63,61%	[22,9 ; 28,7]
Offensivkurs – Wenig Vorwissen	13,18	32,15%	[10,2 ; 16,2]
Offensivkurs – Viel Vorwissen	26	63,41%	[22,9 ; 29,1]

Tabelle 5.3: Ergebnisse des Wissenstests vor Kursdurchführung

5.3.1 Awarenesstest vor Kursbeginn

Im zweiten Teil des Fragebogens, dem Awarenesstest, wurden die Studenten hinsichtlich ihrer Einstellung zu IT-sicherheitsrelevanten Themen befragt. Dabei differierten die Ergebnisse vor der Durchführung der Kurse vor allem bei den Teilnehmern des offensiven Kurses, was aber dadurch erklärbar ist, dass es sich hierbei um eine echt zufällige Verteilung handelte, während bei der Aufteilung nach Vorwissen auf eine ungefähre

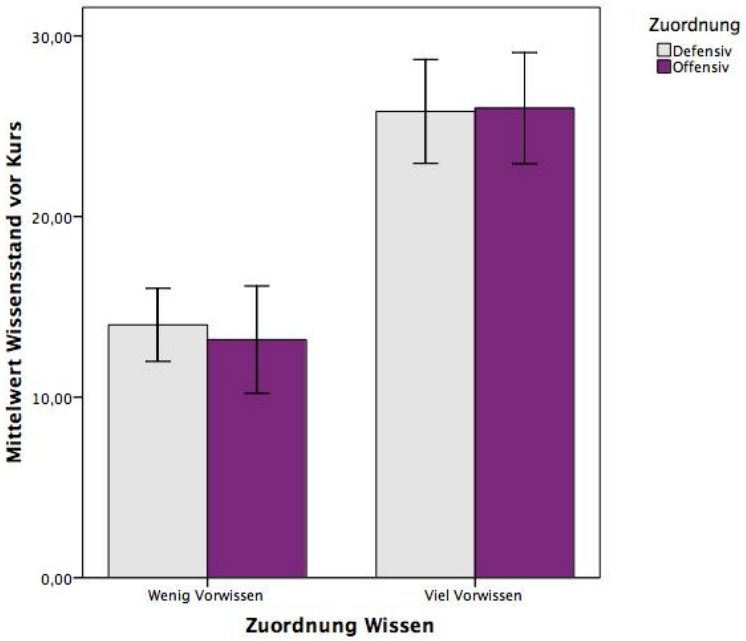


Abbildung 5.4: Ergebnisse des Wissenstests vor Kursdurchführung

Gleichverteilung geachtet wurde. Die Ergebnisse sind in Abbildung 5.5 bzw. Tabelle 5.4 dargestellt.

	Punkte	in %	Konfidenzintervall
Defensivkurs – Wenig Vorwissen	13,8	71,67%	[11,4 ; 16,2]
Defensivkurs – Viel Vorwissen	13,36	70,44%	[10,1 ; 16,7]
Offensivkurs – Wenig Vorwissen	9,73	60,36%	[7,3 ; 12,2]
Offensivkurs – Viel Vorwissen	14,64	74%	[12,5 ; 16,8]

Tabelle 5.4: Ergebnisse des Awarenesstests vor Kursdurchführung

Die Auswertung dieses Fragebogens ergab, dass 85% aller Teilnehmern den Schutz des eigenen Systems als wichtig bzw. sehr wichtig empfinden und sich auch sehr oft in Fachzeitschriften über IT-Sicherheit informieren, aber gerade einmal 30% ihre Daten bzw. E-Mails verschlüsseln, regelmäßig Sicherheitsupdates ihres Betriebssystems ausführen oder Backups ihrer wichtigen Daten erstellen. Diese Diskrepanz zwischen Sicherheitsbewusst-

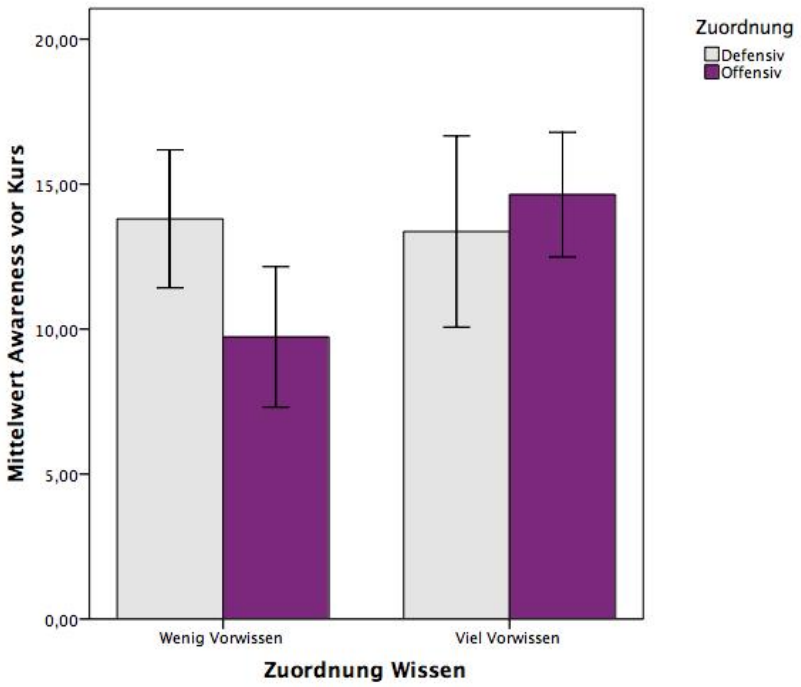


Abbildung 5.5: Ergebnisse des Awarenesstests vor Kursdurchführung

sein und Handlungsbereitschaft ist gerade im Bereich der IT-Sicherheit ein häufig anzutreffendes und viel diskutiertes Problem, was nochmals die Wichtigkeit von Awareness- und Trainingsveranstaltungen, wie sie in Abschnitt 2.2 vorgestellt wurden, untermauert.

5.3.2 Tests am Kursende

Interessant ist die Betrachtung der Ergebnisse, welche die Teilnehmer nach der Kursdurchführung in den beiden Tests erzielt haben. Dazu erhielten die Studenten einen zum Vortest identischen Wissenstest und einen etwas abgewandelten Fragebogen zur Awareness (siehe Anhang B). Bei der Betrachtung der Ergebnisse des Wissenstests in Abbildung 5.6 und Tabelle 5.5 fällt auf, dass neben einer allgemeinen Verbesserung aller Teilnehmer – was zu erwarten war – vor allem die Studenten aus dem offensiven Kurs deutlich bessere Werte erzielt haben, also mehr Antworten korrekt beantworten konnten,

als die Teilnehmer aus dem Defensivkurs.

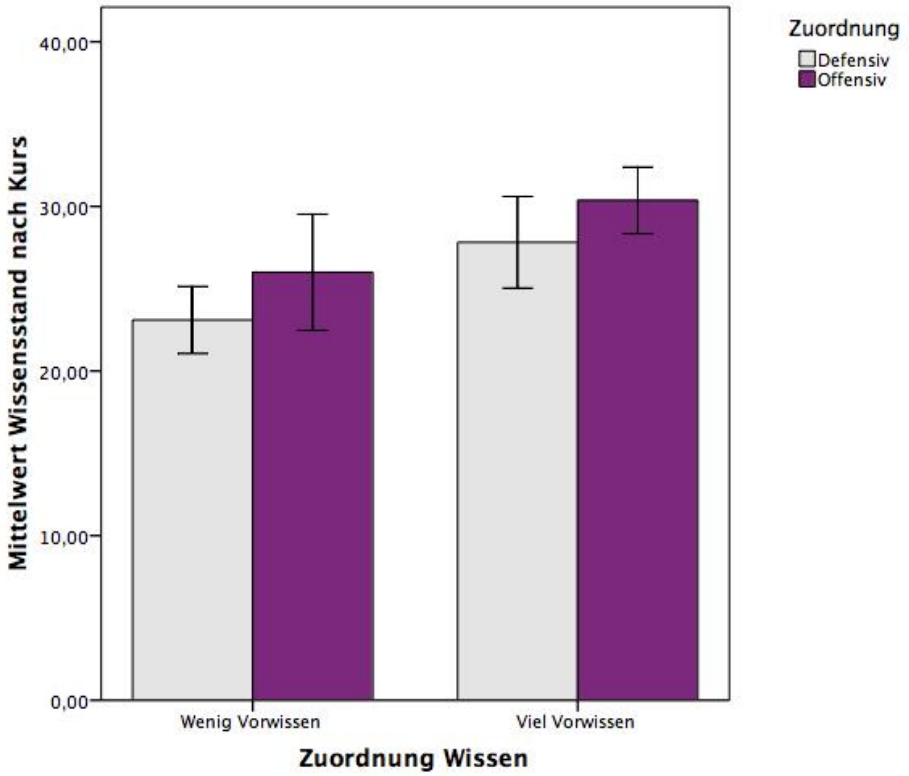


Abbildung 5.6: Ergebnisse des Wissenstests nach Kursdurchführung

Bezüglich der Einstellung der Teilnehmer gegenüber IT-Sicherheit lassen sich hingegen durch einen bloßen Vergleich der Vor- und Nachtest-Ergebnisse keine Aussagen über Einflüsse des offensiven oder defensiven Lehransatzes machen, wie Abbildung 5.7 und Tabelle 5.6 zeigen. Erkennbar ist eine generelle Verbesserung im Vergleich zu den Resultaten vor Kursbeginn, was für eine zunehmende Sensibilisierung der Teilnehmer spricht, jedoch vorauszusehen war.

Doch nicht nur die direkt erzielten Testergebnisse sind für eine Bewertung der beiden Lehransätze relevant, sondern auch die Fortschritte, die die Studenten durch die Teilnahme am Kurs gemacht haben. Daher betrachten die folgenden Abbildungen die relativen Veränderungen der Teilnehmer zu ihren Resultaten vor Beginn der Kurse.

	Punkte	in %	Konfidenzintervall
Defensivkurs - Wenig Vorwissen	23,1	56,34%	[21,1 ; 25,1]
Defensivkurs - Viel Vorwissen	27,82	67,85%	[25,0 ; 30,6]
Offensivkurs - Wenig Vorwissen	26	63,41%	[22,5 ; 29,5]
Offensivkurs - Viel Vorwissen	30,36	74,05%	[28,3 ; 32,4]

Tabelle 5.5: Ergebnisse des Wissenstest nach Kursdurchführung

	Punkte	in %	Konfidenzintervall
Defensivkurs - Wenig Vorwissen	15,7	76,94%	[13,6 ; 17,8]
Defensivkurs - Viel Vorwissen	15,09	75,25%	[12,2 ; 18,0]
Offensivkurs - Wenig Vorwissen	12,73	68,69%	[10,7 ; 14,8]
Offensivkurs - Viel Vorwissen	16,64	79,56%	[14,6 ; 18,6]

Tabelle 5.6: Ergebnisse des Awarenesstests nach Kursdurchführung

In Abbildung 5.8 ist zu sehen, dass die Unterschiede zwischen Defensiv- und Offensivkurs bezüglich des IT-Sicherheitswissens noch deutlicher ausgeprägt sind als in Abbildung 5.6. So konnten die Offensivkursteilnehmer ihre Ergebnisse aus dem Wissenstest um 43,85% verbessern, die Teilnehmer aus dem Defensivkurs lediglich um 27,88%. Das entspricht, wie in Abbildung 5.9 dargestellt, einem um 57,28% besseren Ergebnis als der Defensivkurs.

Auch bezüglich der Awareness der Teilnehmer ist im Gegensatz zu den Resultaten aus Abbildung 5.7 eine Aussage möglich. So scheinen die Studenten des Offensivkurses durch die Erfahrungen mit Angriffstechniken eher bereit zu sein, ihr gegenwärtiges Verhalten im Umgang mit Computern zu überdenken und z.B. mehr auf die Verschlüsselung von E-Mails und Daten zu achten oder häufiger Backups durchzuführen. Wie Abbildung 5.9 zeigt, wurden sie durch die Kursdurchführung um 45,7% stärker sensibilisiert bezüglich potentieller IT-Sicherheitsrisiken als die Teilnehmer des Defensivkurses. Diese Ergebnisse des Awarenesstests werden durch die Auswertung einer weiteren Zusatzfrage untermauert, in welcher die Teilnehmer die Sicherheit ihres eigenen Systems einschätzen sollten. Dabei empfanden 36,4% der Offensivkursteilnehmer ihr eigenes System als unsicher, hingegen nur 14,29% der Defensivkursteilnehmer (siehe Abbildung 5.10).

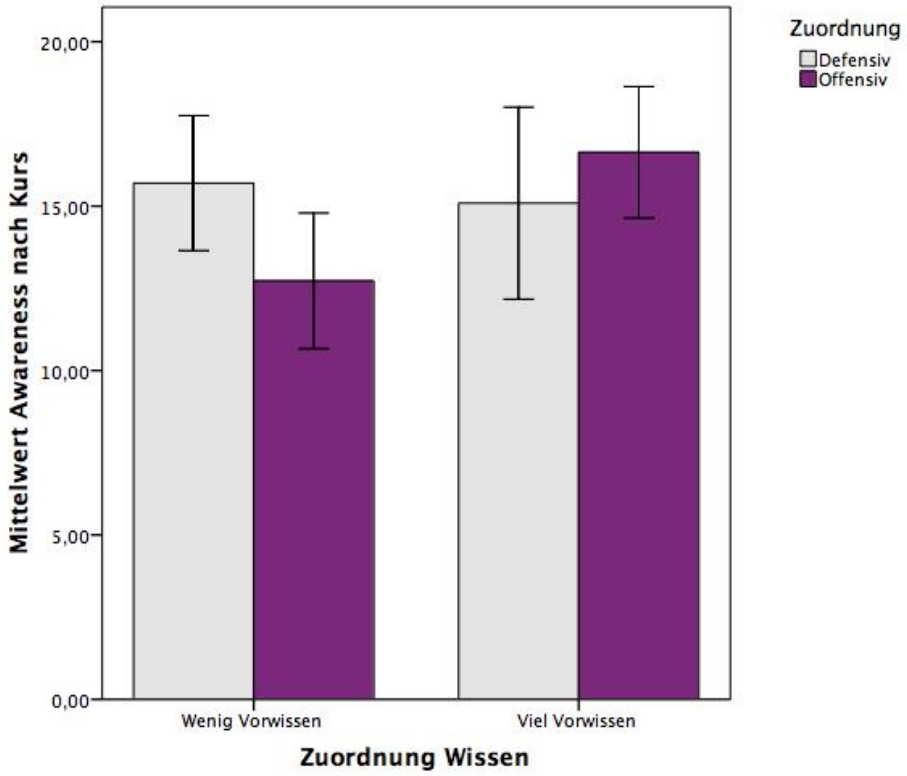
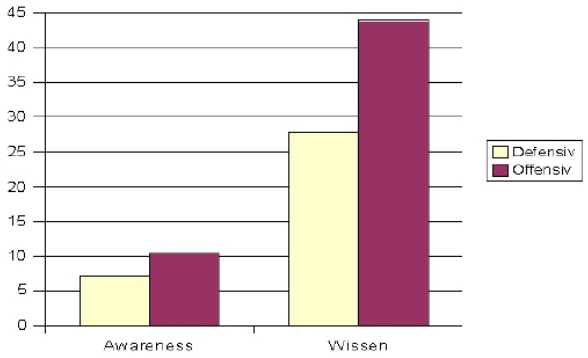


Abbildung 5.7: Ergebnisse des Awarenesstests nach Kursdurchführung

Die Awarenessangaben haben aber, wie bereits angedeutet, nur eine begrenzte Aussagekraft, da der Unterschied zwischen IT-Sicherheitsbewusstsein und sicherem Handeln oft sehr groß ist. Daher können diese Daten nur die Interpretation der Ergebnisse anderer Tests verstärken bzw. abschwächen, alle anderen alleinigen Behauptungen über die Awarenessergebnisse sind mit entsprechender Vorsicht zu behandeln.

5.4 Aussagekraft der Ergebnisse

Die bisher vorgestellten und betrachteten Ergebnisse zeigen zwar, dass die Teilnehmer des Offensivkurses durchweg bessere Ergebnisse erzielt haben als die Studenten der Kontrollgruppe, allerdings muss nun die bereits mehrfach angesprochene Aussagekraft der



	Awarenesstest	Wissenstest
Defensivkurs	+7,09%	+27,88%
Offensivkurs	+10,33%	+43,85%

Abbildung 5.8: Fortschritte der Teilnehmer

Ergebnisse sowohl statistisch, als auch in einer kritischen Reflexion der Studie, untersucht werden.

5.4.1 Signifikanztests

Zur Überprüfung einer Hypothese, d.h. ob sie bestätigt oder widerlegt werden kann, können verschiedene Signifikanztests durchgeführt werden. So können zum einen grafische Analysen zur Bestimmung der Aussagekraft und Interpretationsfähigkeit der Ergebnisse durchgeführt werden, zum anderen auch statistische Tests zur Berechnung der Irrtumswahrscheinlichkeit, die Nullhypothese zugunsten der zu überprüfenden Forschungshypothese zu verwerfen. Zur Erinnerung, die Nullhypothese beschreibt das Gegenteil der Forschungshypothese, in diesem Fall, dass der offensive Lehransatz *nicht* besser ist als der defensive Ansatz.

Für die Wahl des geeigneten Signifikanztests ist die Wahl des Untersuchungsdesigns entscheidend. In Abschnitt 4.1.3 wurde ein zweifaktorieller Untersuchungsplan gewählt, bei welchem die Bedeutung von zwei unabhängigen Variablen auf eine abhängige Variable untersucht wird. Dieser Versuchsplan ermöglicht im Rahmen einer zweifaktoriellen

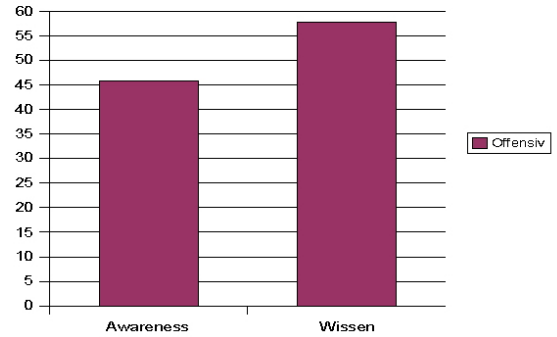


Abbildung 5.9: Relativer Unterschied der Teilnehmerfortschritte

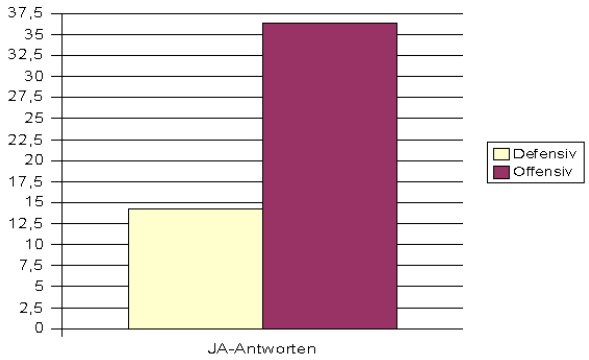


Abbildung 5.10: Teilnehmer, die ihr System als unsicher empfinden

Varianzanalyse (ANOVA) die Überprüfung der beiden Haupteffekte Kurszuordnung und Vorwissenszuordnung sowie die Interaktionseffekte, also die Kombinationswirkung dieser beiden Effekte, auf das Ergebnis.

Die Varianzanalyse wird hier nur auf den Abschlusstest angewendet, da dessen Ergebnisse das entscheidende Kriterium für die Bewertung der beiden Lehransätze darstellen. Daraus ergeben sich drei Hypothesen für die Überprüfung der Haupt- und der Interaktionseffekte, ob sie global interpretiert werden können oder nicht:

1. Die Kurszuordnungen *offensiv* und *defensiv* wirken unabhängig vom Vorwissen unterschiedlich auf die Ergebnisse des Abschlusstests. Eine globale Interpretation

für den Haupteffekt der Kurszuordnung ist z.B., dass der Offensivkurs besser ist als der Defensivkurs (Haupteffekt A).

2. Die Untergruppen *wenig* oder *viel* Vorwissen haben in Bezug auf die Kurszuordnung unterschiedliche Auswirkungen auf die Ergebnisse des Abschlusstests. Eine globale Interpretation in Bezug auf das Vorwissen könnte sein, dass Studenten mit viel Vorwissen bessere Ergebnisse im Abschlusstest erzielen als Studenten mit wenig Vorwissen (Haupteffekt B).

3. Es gibt eine differentielle Wirkung zwischen der Kurszuordnung und der Vorwissenszuordnung, z.B. dass der Offensivkurs besser ist für Studenten mit viel Vorwissen (Interaktion A×B).

Diese drei Hypothesen lassen sich nun sowohl grafisch als auch statistisch überprüfen, wobei der statistische Ansatz zwar komplizierter ist, aber dafür auch viel genauere Aussagen über die Interpretation und Signifikanz der Ergebnisse geben kann. Dennoch soll hier zuerst die grafische Ermittlung durchgeführt werden, um erste Hinweise auf die Aussagekraft der Ergebnisse zu erhalten, bevor im Anschluss die exakten statistischen Werte berechnet werden.

5.4.2 Grafische Analyse der Signifikanz

Für eine grafische Signifikanzanalyse der Haupt- und Interaktionseffekte werden die Zellenmittelwerte, nämlich wie viele Punkte jede der vier Untergruppen im Durchschnitt erreicht hat, sowie die Randmittelwerte aus den Spalten bzw. Zeilen berechnet, wie sie in Tabelle 5.7 dargestellt sind.

	Defensiver Kurs	Offensiver Kurs	Mittelwert B
Wenig Vorwissen	5,45	5,41	**5,43**
Viel Vorwissen	5,96	6,82	**6,39**
Mittelwert A	**5,7**	**6,11**	**5,91**

Tabelle 5.7: Zweifaktorielle ANOVA des Abschlusstests

Ein erster rechnerischer Ansatz für die Überprüfung eines signifikanten Haupteffekts besteht in der Subtraktion der Spalten- bzw. Zeilenmittelwerte. Je größer die Differenz zwischen den Mittelwerten ist, desto eher spricht dies für einen signifikanten Haupteffekt. Aus der obigen Tabelle ergibt sich für den Haupteffekt A, die Kurszuordnung, eine Differenz von

$6,11 - 5,7 = 0,41$

und für den Haupteffekt B, die Untergruppenzuordnung, eine Differenz von

$6,39 - 5,43 = 0,96.$

Eine genaue graphische Analyse erfolgt durch die Betrachtung von Interaktionsdiagrammen, in welchen die Zellenmittelwerte aus Tabelle 5.7 betrachtet werden. Hierfür werden auf der y-Achse die Ergebnisse aus dem Abschlusstest eingetragen und auf der x-Achse der jeweils betrachtete Haupteffekt (Kurszuordnung oder Vorwissenszuordnung). Es resultieren also zwei Diagramme, wie sie in Abbildung 5.11 dargestellt sind.

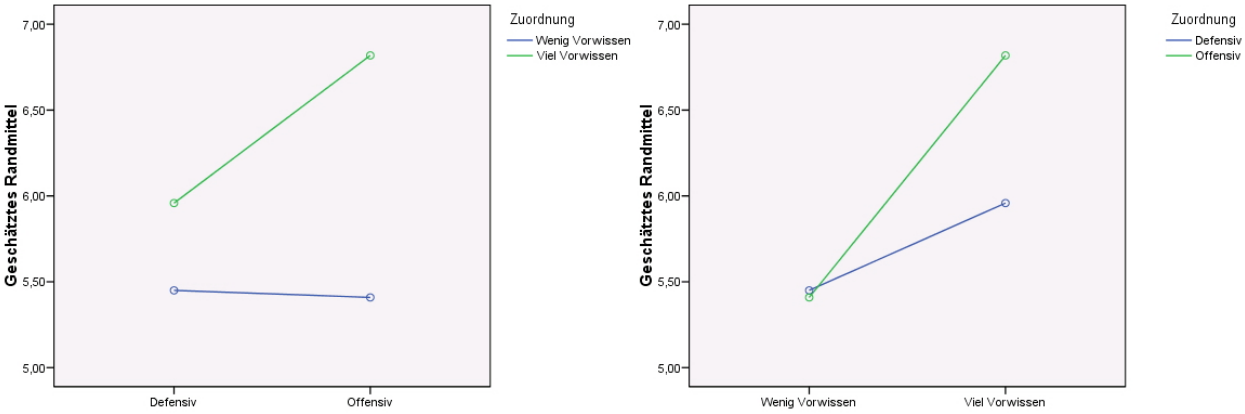


Abbildung 5.11: Interaktionsdiagramme für die Ergebnisse des Abschlusstests

Entscheidend für die Interpretationsfähigkeit der beiden Haupteffekte ist der Verlauf der Geraden in beiden Diagrammen, anhand dessen der so genannte Interaktionstyp bestimmt werden kann. Verlaufen die Geraden in beiden Diagrammen gleichsinnig, also in beiden Darstellungen entweder steigend oder fallend, spricht man von einer *ordinalen Interaktion*. Dieser Interaktionstyp erlaubt eine globale Interpretation beider Haupteffekte. Verlaufen die Geraden jedoch nicht gleichsinnig, wie es in den hier dargestellten

Interaktionsdiagrammen der Fall ist, handelt es sich um eine *hybride Interaktion*, was bedeutet, dass nur ein Haupteffekt global interpretiert werden kann.

Dass es sich hierbei um den Haupteffekt B (Vorwissen) handelt, wird neben der obigen Berechnung auch durch eine genauere Betrachtung der Diagramme bestätigt. Im linken Diagramm weist die Gruppe mit wenig Vorwissen niedrigere Werte auf als die Gruppe mit viel Vorwissen. Im Vergleich dazu liegen aber im rechten Diagramm nicht alle Werte des Defensivkurses unter denen des Offensivkurses, weswegen der Effekt der Kurszuordnung nicht global interpretiert werden kann.

Dem Verlauf der Geraden nach zu urteilen scheint auch eine differentielle Wirkung der Faktoren Vorwissen und Lehransatz zu bestehen. So fallen im linken Diagramm die Werte für die Untergruppe mit wenig Vorwissen von Defensiv zu Offensiv, und im rechten Diagramm liegen die Ergebnisse der Defensivgruppe für wenig Vorwissen sogar über den Ergebnissen des Offensivkurses. Es kann also sein, dass ein defensiver Kurs besser geeignet ist für Personen mit wenig IT-Sicherheitswissen, und der Offensivkurs besser für Personen mit viel Vorwissen.

Wie bereits beschrieben ist eine Signifikanzprüfung anhand von Interaktionsdiagrammen zwar sehr anschaulich und leicht verständlich, aber nicht so genau wie die exakte Berechnung des Signifikanzwertes, was durch den folgenden Abschnitt ersichtlich wird.

5.4.3 Statistische Ermittlung der Irrtumswahrscheinlichkeit

Die eben durchgeführte zweifaktorielle ANOVA kann auch zur Bestimmung eines konkreten Signifikanzwertes verwendet werden, des so genannten p-Wertes. Der *p-Wert* gibt die Wahrscheinlichkeit an, mit der die Ergebnisse zufällig entstanden sein können. Liegt der p-Wert unter einem zuvor festgelegten Signifikanzniveau – meist 5% – spricht man von einem signifikanten Ergebnis. Diese Wahrscheinlichkeit wird daher auch als die Irrtumswahrscheinlichkeit bezeichnet, dass eine Nullhypothese fälschlicherweise verworfen wird.

Die Berechnung des p-Wertes ohne ein entsprechendes Statistikprogramm ist sehr zeitintensiv, komplex und erfordert ein hohes Maß an statistischen Kenntnissen, weswegen für ein tiefergehendes Interesse der Berechnung auf [Hartung u. a., 2005] verwiesen wird.

Abbildung 5.12 zeigt die Ergebnisse der p-Wert-Berechnung mittels SPSS.

Abhängige Variable: Abschluss-Test

Quelle	Quadratsumme vom Typ III	df	Mittel der Quadrate	F	Signifikanz
Kurs	1,837	1	1,837	,404	,529
Untergruppe	10,068	1	10,068	2,213	,145
Kurs * Untergruppe	2,222	1	2,222	,488	,489
Fehler	182,000	40	4,550		

Abbildung 5.12: Berechnung des p-Wertes für den Abschlusstest

Die relevanten Ergebnisse aus dieser Abbildung sind die Signifikanzwerte (p-Werte) der Spalte ganz rechts. Es ist zu erkennen, dass alle Werte weit über dem festgelegten Signifikanzniveau von 0,05 liegen und somit keine Aussage bezüglich der Kurszuordnung, der Untergruppenzuordnung nach Vorwissen oder der Interaktion zwischen diesen beiden als signifikant angesehen werden kann. So liegen die Irrtumswahrscheinlichkeiten für die Behauptungen, dass

- der offensive Lehransatz besser ist als der defensive Ansatz bei 52,9%
- das IT-Sicherheitsverständnis bei Studenten mit viel Vorwissen ausgeprägter ist als bei Studenten mit wenig Vorwissen bei 14,5%
- der Defensivkurs besser geeignet ist für Gruppen mit wenig Vorwissen, und der Offensivkurs besser für Gruppen mit viel Vorwissen bei 48,9%

Die Ergebnisse bedeuten nicht, dass diese Aussagen damit widerlegt sind oder das Modell der Studie schlecht war, sondern nur, dass anhand des erhobenen Datenmaterials keine signifikanten Unterschiede zwischen den Kursergebnissen festzustellen sind und daher mehr Daten für eine höhere Aussagekraft erforderlich sind, z.B. durch eine Wiederholung der Studie. Außerdem sollte überprüft werden, ob die Stichprobengröße größer bemessen werden sollte.

5.5 Kritische Reflexion der Studie

Der letzte Abschnitt zeigte bereits, dass die Ergebnisse der Studie nicht signifikant sind. Eine kritische Reflexion der Studie soll nun hinterfragen, ob die mangelnde Signifikanz auf zu wenig Datenmaterial und eine zu kleine Stichprobe zurückzuführen ist, oder ob auch systematische Fehler in der Studie, z.B. in der Planung oder Durchführung, zu ungewollten Verzerrungen der Ergebnisse geführt haben können. Dazu sind vier relevante Aspekte näher zu betrachten:

- Versuchsleitereffekte (siehe Abschnitt 5.5.1)
- Technische Probleme (siehe Abschnitt 5.5.2)
- Inhaltliche Probleme (siehe Abschnitt 5.5.3)
- Feedback der Kursteilnehmer (siehe Abschnitt 5.5.4)

5.5.1 Versuchsleitereffekte

Der Versuchsleitereffekt ist eine nicht zu vernachlässigende Fehlerquelle bei der Durchführung von empirischen Studien, welcher zu erheblichen Verzerrungen der Untersuchungsergebnisse führen kann. Man unterscheidet zwischen Effekten durch Erwartungshaltungen des Versuchsleiters oder Verzerrungen durch Lern- und Gewöhnungseffekte. Das Problem einer unbeabsichtigten Beeinflussung der Ergebnisse durch entsprechende Erwartungshaltungen wird auch als Rosenthal-Effekt [Bortz u. Döring, 2006] bezeichnet. Dieser Effekt besagt, dass sich die Leistungen von Versuchspersonen bereits durch die Einstellung des Versuchsleiters in eine erwartete Richtung entwickeln können, z.B. wenn der Versuchsleiter dieser Studie mit der Überzeugung die Kompaktkurse durchführt, dass die offensive Ausrichtung zu einem höheren IT-Sicherheitsverständnis führt als die defensive Lehre, und so durch unbewusstes Verhalten die Ergebnisse beeinflusst. Eine zweite mögliche Gefahr der Ergebnisverzerrung besteht durch Lern- und Gewöhnungseffekte des Versuchsleiters, z.B. die Reaktionen auf Nachfragen, Korrektur von Fehlern auf Fo-

lien bzw. im Arbeitsmaterial zwischen den Kursen oder allgemeine Verbesserungen im Präsentationsstil.

Um dem Versuchsleitereffekt entgegenzuwirken, sollten im Idealfall für eine Gewährleistung der Validität der Untersuchungsergebnisse Forscher und Versuchsleiter unterschiedliche Personen sein, womit der Versuchsleiter völlig unbeeinflusst und nicht in die Studie involviert ist. Diese Fehlerquelle ließ sich nicht vermeiden, da in der durchgeführten Studie Dozenten und Versuchsleiter identisch waren. Jedoch kann durch die Vermittlung der Lehrinhalte durch drei verschiedene Dozenten – wobei jeder Dozent in beiden Kursen dieselben Module geleitet und vorgetragen hat – eine Verzerrung durch den Rosenthal-Effekt mit hoher Wahrscheinlichkeit ausgeschlossen werden. Zukünftige Studien könnten versuchen, Dozenten zu finden, die nicht über die Hypothese informiert sind oder die Wissensvermittlung standardisiert über ein E-Learning-Tool zu gestalten.

Nicht gänzlich auszuschließen waren allerdings Verzerrungen durch Lern- und Gewöhnungseffekte der Versuchsleiter. So wurde zwar darauf geachtet, die Teilnehmer in beiden Kursen möglichst gleich zu behandeln, allerdings wurden im Offensivkurs die Vorträge zur Theorie aufgrund der Wiederholungseffekte insgesamt besser und sicherer präsentiert. Ob dies eine Auswirkung auf den Lernerfolg und somit die Ergebnisse der Teilnehmer hatte, ist zu diesem Zeitpunkt nicht zu beantworten, sondern erst, wenn Vergleichswerte aus einer Wiederholung der Studie vorliegen.

5.5.2 Technische Probleme

Sinkende Motivation und schlechte Arbeitsbedingungen, hervorgerufen durch Soft- oder Hardwareprobleme waren ein großes Problem im defensiven Kurs, der zuerst stattfand. So kam es durch die Kombination von Microsoft Virtual PC mit Linux zu unerwarteten Systemabstürzen und Speicherplatzproblemen auf der physischen Festplatte, da die virtuellen Festplatten stetig in ihrer Größe wuchsen, auch wenn Daten gelöscht wurden. Zusätzlich war der zur Verfügung stehende Speicherplatz auf der physischen Festplatte relativ klein.

Die technischen Probleme führten vor allem zu Verzögerungen im Tagesablauf, wodurch Vorträge verspätet begannen, die praktischen Übungen kürzer ausfielen oder die

Besprechung der Musterlösungen am Ende der Module teilweise entfallen musste. Im offensiven Kurs bestanden zwar auch Probleme durch gelegentliche Systemabstürze, jedoch konnten die Versuchsleiter durch die gemachten Erfahrung aus der Vorwoche den Speicherplatzproblemen frühzeitig entgegenwirken, wodurch es keine Verzögerungen im Kursablauf gab und auch genügend Zeit für die Besprechung der Übungen blieb.

5.5.3 Inhaltliche Probleme

Wie schon in Abschnitt 4.2 deutlich wurde, war es teilweise schwer, eine sinnvolle und eindeutige Abgrenzung zwischen offensiven und defensiven Übungen zu finden, da viele Angreifertechniken ebenfalls von Administratoren zur Erhöhung der IT-Sicherheit angewendet werden. Dadurch gab es viele Gemeinsamkeiten zwischen den Übungen, weswegen kein Kurs als klar offensiv oder defensiv bezeichnet werden kann. Für eine Wiederholung der Studie kommt daher auch eine Anpassung der theoretischen Inhalte in Betracht, welche bisher identisch gehalten wurden. Eine Auswahl geeigneter offensiver bzw. defensiver Kursthemen kann aus [Mertens, 2007] erstellt werden. Es muss jedoch sichergestellt werden, dass die Teilnehmer beider Kurse die nötigen Grundvoraussetzungen vermittelt bekommen, um den praktischen Abschlusstest bewältigen zu können.

5.5.4 Feedback der Kursteilnehmer

Die Kursteilnehmer erhielten zum Abschluss des Kurses einen Kursbewertungsbogen, in welchem sie anonym die Durchführung des Kurses bewerten und ihre persönliche Meinung niederschreiben konnten, gefolgt von einer offenen Diskussion mit allen Teilnehmern. Sowohl im Bewertungsbogen, als auch in der Diskussionsrunde äußerten sich fast alle Teilnehmer trotz technischer Probleme sehr zufrieden über die Kursdurchführung und bestätigten, einen guten und umfassenden Einblick in das Gebiet der IT-Sicherheit erhalten zu haben. Vor allem die Kombination von Theorie und direktem Anschluss der praktischen Übungen, um das Gelernte anwenden zu können, wurde äußerst positiv aufgenommen und hat den Teilnehmern auch sichtlich Spaß gemacht.

Ein Hauptkritikpunkt war jedoch die mangelnde Zeit. So wurde häufig der Wunsch

geäußert, die Kursdauer auf mindestens 5 Tage zu erhöhen und auf die Module „Einführung“ und „Firewalls“ zu verzichten, um die anderen Module dafür länger und detaillierter in der Theorie behandeln zu können.

Ein weiterer geäußerter Kritikpunkt war, dass Teilnehmer mit sehr geringen Vorkenntnissen in Linux und der Programmiersprache C die Kurse nicht gut bewältigen konnten und schnell mit den Themen überfordert waren, auch wenn diese Gebiete in der Einführungsveranstaltung behandelt wurden. Daher sollten in einer erneuten Kursausschreibung mehr Vorkenntnisse gefordert werden.

5.6 Zusammenfassung und Ausblick

Die Auswertung der Ergebnisse führt unter Berücksichtigung der beschriebenen Probleme und Resultate der Signifikanztests zu der Schlussfolgerung, dass die zu Beginn der Studie aufgestellte Forschungshypothese nicht bestätigt werden kann. Sie ist allerdings damit auch nicht widerlegt, sondern muss durch mehr Datenmaterial aus einer Wiederholung der Studie weiter geprüft werden. Es kann jedoch nicht ausgeschlossen werden, dass auch Fehler in der Studienplanung und -durchführung, wie technische Probleme oder die unzureichenden Abgrenzungen zwischen dem offensiven und defensiven Ansatz, zu einem Verlust der Signifikanz geführt haben.

6 Fazit und Ausblick

In dieser Arbeit wurde der Stand der universitären Ausbildung in Informationssicherheit in Deutschland und der Welt vorgestellt. Es zeigte sich, dass sich die untersuchten Einführungsveranstaltungen in drei Klassen gruppieren lassen: eine konservative, in der Kryptografie einen großen Anteil hat, eine innovative, in der aktuelle Themen der IT-Sicherheit behandelt werden, sowie eine, in der alle Themen in etwa gleich behandelt werden. Es wurde festgestellt, dass die Dozenten von IT-Sicherheitsvorlesungen an deutschen Universitäten jünger sind als ihre Kollegen an den betrachteten internationalen Universitäten und dass es keinen signifikanten Zusammenhang zwischen dem Alter des Dozenten und den von ihm gelehrten Themen gibt. Es ließ sich lediglich die Tendenz erkennen, dass jüngere Dozenten aktuellere Themen lehren. Vorlesungen, die Teil der konservativen Klasse sind, sind bereits seit längerer Zeit in den Lehrplänen enthalten, als die der innovativen Klasse. Bei der im innovativen Cluster verwendeten Literatur handelt es sich um Bücher zu grundlegenden Themen im Bereich IT-Sicherheit und solchen zu Verteidigungsmaßnahmen und Angriffsmethoden. Die Vorlesungen des ausgewogenen Cluster basieren sowohl auf Büchern über Kryptografie als auch auf wissenschaftlichen Veröffentlichungen und Webquellen. Zudem wird Literatur verwendet, die Verteidigungsmaßnahmen und Angriffsmethoden enthalten. Die im konservativen Cluster verwendete Literatur deckt schwerpunktmäßig den Bereich Kryptografie ab. Außerdem wurde festgestellt, dass in den Veranstaltungen an internationalen Universitäten aktuelle Literatur wie Konferenzveröffentlichungen und Webquellen den größten Anteil ausmachen, während diese an deutschen Universitäten eine geringe Rolle spielen.

Hauptteil der Arbeit war jedoch die Bewertung des offensiven Ansatzes in der Ausbildung in IT-Sicherheit an Universitäten. Als *offensiv* werden Techniken bezeichnet,

die darauf abzielen, „etwas kaputt zu machen“. Eine konkretisierende Definition und Beispiele für offensive Methoden in Lehrveranstaltungen und Lehrformen wurden in Kapitel 2 vorgestellt. Zum Zweck der Bewertung wurde eine empirische Studie geplant und durchgeführt. Im Rahmen der Studie wurden zwei Kurse zu IT-Sicherheit entworfen und durchgeführt. Der eine Kurs entspricht den Inhalten des innovativen Clusters (stellvertretend für die offensive Ausbildung), der andere denen des konservativen Clusters (stellvertretend für die defensive Ausbildung). Beide Kurse haben praktische Anteile. Die Studie überprüfte die Auswirkung der Kurse auf das IT-Sicherheitswissen und -verständnis der Teilnehmer. Die Ergebnisse basieren auf einem schriftlichen Test zu Einstellung und Wissen, der zu mehreren Zeitpunkten durchgeführt wurde, sowie einem praktischen Test am Ende des Kurses, der die Probanden vor die Aufgabe stellte, eine Betriebssysteminstallation auf Schwachstellen zu überprüfen und in einen sicheren Zustand zu überführen. Ausgewertet wurde, wie viele Sicherheitsprobleme die Probanden auf dem präparierten System fanden, wie sie die Probleme lösten und welche globale Strategie sie verfolgten. Die Ergebnisse der Studie werden in der Übersicht vorgestellt.

Die Teilnehmer des Offensivkurses haben ein etwas besseres Ergebnis erzielt, als die des Defensivkurses. Allerdings spricht die Überprüfung der Konfidenz dafür, dass es sich um ein nicht aussagekräftiges Ergebnis handelt. Bei Unterscheidung nach Vorwissen wird das bessere Abschneiden des Offensivkurses deutlicher – allerdings nur für die Subgruppe mit viel Vorwissen. Und auch hier ließ sich kein signifikanter Unterschied feststellen.

Die Teilnehmer des Offensivkurses zeigen jedoch eine stärkere Sensibilisierung und haben einen größeren Wissenszuwachs. Im Vergleich der Tests am Kursanfang und am Ende konnten sie ihre Ergebnisse sowohl im Wissens- als auch im Awarenesstest stärker steigern als der Defensivkurs.

Außerdem ließ sich feststellen, dass die Teilnehmer mit mehr Vorwissen bessere Ergebnisse erzielt und die Teilnehmer insgesamt einen Wissenszuwachs erreicht haben. Dies war zu erwarten, spricht aber für die Vermittlung der Inhalte sowie die Konzeption und Durchführung der Studie.

Derzeit ist also – basierend auf den in der Studiendurchführung gesammelten empirischen Daten – keine definitive Aussage möglich, ob der offensive Ansatz generell besser

ist. Die offensive Ausbildung erweist sich aber als ein interessanter, mittlerweile weit praktizierter Ansatz in der universitären Ausbildung. Aus der Erfahrung des Autoren hat die sie gewisse Vorteile gegenüber einer rein defensiven Ausbildung. Denn die offensive Ausbildung führt zu einem besseren Verständnis von Sicherheitsproblemen und -lücken. Der offensive Ansatz ist motivierender, da zum einen die Unsicherheit eines Systems einfacher zu zeigen ist – nämlich durch eine einzige Sicherheitslücke – als die Sicherheit des Systems. Zum anderen sind motivierendere Lehrformen möglich, die spielerische Aspekte beinhalten, wie die vorgestellten CTF-Wettbewerbe und Wargames. Ein weiterer, wichtiger Aspekt ist die Angreiferperspektive und damit das Kennenlernen von Angreiferdenken und -sicht.

Wie wichtig eine offensive Herangehensweise ist, lässt sich an Bereichen erkennen, in denen offensive Techniken bereits länger verwendet werden. So gibt es erst sehr gute kryptografische Algorithmen, seitdem öffentlich daran geforscht wird und auch Kryptoanalyse, d.h. Angriffe auf die Verfahren, angewandt wird. Aber auch in der produzierenden Industrie werden bereits seit langem offensive Methoden zur Fehlersuche und Schwachstellenerkennung eingesetzt. Beispielsweise werden im Automobilbereich Crashtests zur Erhöhung der Sicherheit genutzt.

Nichtsdestotrotz muss auch eine offensiv ausgebildete Person gängige Abwehrmaßnahmen kennen und wissen, wie und wo diese anzuwenden sind. So bietet die offensive Ausbildung auf jeden Fall eine gute Ergänzung zur defensiven Ausbildung, d.h. eine Kombination beider Ansätze. Dies wird gestützt durch die Untersuchung und Klassifikation von IT-Sicherheitsveranstaltungen [Mertens, 2007]: hier lag der größte Teil der untersuchten Veranstaltungen im ausgewogenen Cluster, d.h. es handelt sich um eine Mischung von offensiven und defensiven Methoden.

Es wird jedoch immer Bedarf bestehen an einer kleine Zahl offensiv ausgebildeter Experten, z.B. zur Landesverteidigung (Stichwort *Cyberwar*), für die Nachrichtendienste oder für die Polizei.

6.1 Weitere Arbeiten

Aus den Erfahrung der ersten Studiendurchführung entstand eine zweite Studiendurchführung, die im Jahr 2008 zuerst an der Universität Mannheim (vom 15.–17. und 22.–24. Januar) und danach erneut an der RWTH Aachen (vom 26.–28. Februar und 4.–6. März) stattfand. Aufgrund von Änderungen im Aufbau sind die dabei erhobenen empirischen Daten nicht mit der Durchführung von 2007 vergleichbar. Da die Auswertung noch nicht abgeschlossen ist, werden die Ergebnisse in einer separaten Veröffentlichung publiziert. Im folgenden werden aber Informationen zum Aufbau der neuen Studiendurchführung gegeben.

Die Anmeldung erfolgte diesmal nicht per E-Mail sondern online über eine Webanwendung, in der die Bewerber einen Zugang einrichteten und ihre Daten eintrugen. Anzugeben waren die gleichen Daten wie im Jahr 2007. Diese Webanwendung wurde im Kurs für den Up- und Download der Kursunterlagen verwendet.

Im Gegensatz zur ersten Studiendurchführung war der Vortest – also der erste Fragebogen – im Jahr 2008 kein Bestandteil des Bewerbungsverfahrens sondern wurde erst mit Kursbeginn erhoben. Durch die zufällige Aufteilung der Teilnehmer auf die beiden Gruppen sollte sich aber eine Gleichverteilung ergeben (was auch in etwa der Fall war). Der Vorteil dieser Lösung ist, dass der Fragebogen in kontrollierter Umgebung und unter den gleichen Bedingungen wie der Fragebogen am Ende des Kurses ausgefüllt wird. Zusätzlich wurde einige Wochen nach Kursende ein weiterer Fragebogen an die Teilnehmer verschickt. Dabei handelt es sich um einen so genannten „Retest“ (siehe Abschnitt 3.1.7). Um die Rücklaufquote dieses Fragebogens zu erhöhen wurde eine garantierte Belohnung (Schokolade) und ein Gewinn (Verlosung eines USB-Speichersticks) in Aussicht gestellt. Tatsächlich beantworteten fast alle Teilnehmer diesen dritten Fragebogen.

Die Bestandteile des Fragebogens – der Awareness und der Wissenstest – waren überarbeitet worden. Einige Fragen wurden weggelassen, dafür andere hinzugefügt und viele umformuliert. Insbesondere das Beantwortungsschema wurde angepasst. So gab es bei den meisten Fragen des Awarenesstests eine Skala von 1–5, die entsprechend der Frage bezeichnet wurde (meist „trifft nicht zu“ bis „trifft voll und ganz zu“). Im Wissenstest

hatte jede Frage vier Antwortalternativen, von denen eine bis vier korrekt sein konnten. Um Raten zu erschweren gab es bei falschen Antworten Punktabzüge. Der Fragebogen enthielt zwei neue Teile: rechtliche Fragen (z.B. „Ich darf fremde unverschlüsselte WLAN-Netzwerke nutzen“) und Fragen zur Legalität (z.B. ob man ohne Auftrag eine Website auf Sicherheitslücken untersuchen darf). Um eine ehrlichere Beantwortung des Fragebogens zu erreichen – insbesondere der Fragen zu Legalität bzw. Illegalität – wurden die Antworten anonym erfasst. Dafür mussten die Teilnehmer auf jedem Fragebogen einen Code angeben, der sich aus bestimmten Buchstaben bzw. Ziffern von persönlichen Daten zusammensetzte (u.a. dem Geburtsdatum und dem Vornamen der Mutter). Durch diesen Code konnten die drei Fragebogen für die Auswertung einander zugeordnet werden, ohne dass Rückschluss auf eine Person möglich ist. Der Fragebogen hatte keine Papierform sondern wurde online ausgefüllt, was den Aufwand für die Erfassung der Antworten und damit auch für die Auswertung reduzierte.

Die Kursinhalte wurden nach den Erfahrungen aus dem Jahr 2007 überarbeitet. Zum einen sollte eine bessere Differenzierung von offensiv und defensiv erreicht werden. Dafür wurden z.B. Passwort-Cracker als offensiv eingeordnet (obwohl diese auch von Administratoren zur Erhöhung der Sicherheit eingesetzt werden) und nur im Offensivkurs eingesetzt; im Defensivkurs wurden stattdessen Passwort-Richtlinien ausführlicher behandelt. Zum anderen wurden Aufgaben aus den Übungen, bei deren Bearbeitung es Probleme gegeben hatte, angepasst. So waren das Erstellen eines Regelwerks für die Firewall (*iptables*) und die Konfiguration des Netzwerks beim ersten Start des virtuellen Rechners sehr zeitaufwändig gewesen. Stattdessen mussten nur noch einzelne Regeln erstellt und andere korrigiert werden bzw. nur noch die IP-Adresse eingetragen werden. Berücksichtigt wurden auch Anmerkungen aus den Feedbackrunden der vergangenen Durchführung, allerdings wurden keine Änderungen an den Themen und der zeitlichen Struktur vorgenommen. Unter anderem war von den Kursteilnehmern das Thema WLAN-Sicherheit gewünscht worden statt eines – nach Ansicht der Teilnehmer weniger interessanten – Themas wie Firewalls. Aber der Verzicht auf ein klassisch defensives Thema wie Firewalls hätte den Verzicht auf eine Differenzierungsmöglichkeit bedeutet.

Als Kurssystem wurde wieder ein Debian-basierendes Linux-System gewählt, als Vir-

tualisierer konnte jedoch diesmal VMware verwendet werden (da in den genutzten PC-Räumen der VMware-Player installiert war).

Für den Abschlusstest wurde ein Linux-System mit grafischer Oberfläche verwendet. Aus diesem Grund wurde kein Keylogger verwendet, weil nicht nachvollziehbar ist, in welchem Fenster die Eingabe stattfand. Zur Protokollierung wurde stattdessen genutzt: die Protokolldateien des Systems (diese wurden über Netzwerk zentral auf dem Kursserver gespeichert) und die Ausgabe eines Shell-Skripts, das alle fünf Minuten protokollierte, welche Dateien sich geändert hatten und welche Ports geöffnet sind. In Mannheim wurde zusätzlich der Bildschirminhalt aufgezeichnet (in Aachen war dies wegen technischer Probleme nicht möglich). Mit der Videoaufzeichnung war eine gute, aber auch sehr zeitintensive Auswertung des Vorgehens möglich.

6.2 Ausblick

Zur weiteren Überprüfung des Beitrags der offensiven Lehre in IT-Sicherheit an Universitäten müssen mehr empirische Daten gesammelt werden, d.h. weitere Studiendurchführungen erfolgen. Aus den Erfahrungen der bisherigen Durchführungen können Verbesserungen erfolgen, um die Studie weiterzuentwickeln. Auch nach der zweiten Studiendurchführung sind weitere Verbesserungen denkbar. Dabei handelt es sich insbesondere um eine (noch) bessere Trennung von offensiven und defensiven Inhalten in den Kursen.

Diesbezüglich sind die folgenden Maßnahmen denkbar:

- Firewall-Modul
  - defensiv: Regelwerk selbst erstellen
  - offensiv: Regelwerk testen und umgehen (*Firewall evasion*)
- Softwaresicherheit-Modul
  - defensiv: Schwachstellen automatisiert (mit Tools) finden und beheben
  - offensiv: Schwachstellen manuell suchen und ausnutzen

- Webanwendungssicherheit-Modul
  - defensiv: sichere Konfiguration eines Webservers, sichere Programmierung von Webanwendungen
  - offensiv: Schwachstellen suchen und ausnutzen

Unterstützend ist eine Überarbeitung der Themen denkbar, konkret die Einführung eines Themenblocks Kryptografie, da sich in der Untersuchung von IT-Sicherheitsveranstaltungen [Mertens, 2007] zeigte, dass dies charakteristisch für defensiv orientierte Kurse ist. Dadurch erhöht sich möglicherweise die Differenzierung der Ansätze. Mögliche Inhalte dieses Krypto-Moduls sind: kryptografische Methoden und Verfahren, Anwendung der Verfahren und Angriffe darauf, Kryptoanalyse (d.h. Angriffe auf kryptografische Verfahren) sowie praktische Anwendungen wie E-Mail-Verschlüsselung. Allerdings ist zu prüfen, ob für dieses Thema passende Inhalte für beide Kurstypen entworfen werden können und diese in den Tests – speziell dem Abschlusstest – entsprechend umgesetzt werden können.

Des weiteren sollte eine bessere Abstimmung der praktischen Inhalte des Offensiv- und des Defensivkurses mit dem Abschlusstest angestrebt werden, d.h. Aufgaben, die mit dem erworbenen Wissen aus den unterschiedlichen Kursen, nämlich Angreifer- oder Verteidigerwissen, gleichermaßen lösbar sind. Ein Beispiel: Ein Teilnehmer des Offensivkurses hat gelernt, dass ein Angreifer Spuren des Angriffs in den Log-Dateien löscht, damit der Angriff nicht entdeckt wird; der Teilnehmer des Defensivkurses hat gelernt, dass die Log-Dateien wichtige Informationen über das System liefern. Aus diesem Wissen sollten beide folgern, dass Log-Dateien nicht für alle Benutzer des Systems schreibbar sein sollten. Für eine optimierte Auswertbarkeit des Abschlusstests sind verbesserte Protokollierungs- sowie Auswertungsmethoden nötig. Um die Beschränkung des Abschlusstests auf Systemadministratorwissen aufzuheben ist eine Erweiterung auf weitere IT-Sicherheitsthemen (z.B. Sicherheitsmanagement) denkbar.

Eine Einschränkung könnten Lehrveranstaltungen mit offensiven Inhalten jedoch durch das Ende Mai 2007 im Bundesrat verabschiedete Strafrechtsänderungsgesetz zur Bekämpfung der Computerkriminalität (§202c StGB, der so genannte „Hacker-Paragraf“) erfahren. Das Gesetz stellt das Herstellen, Verbreiten oder Besitzen von so genannten

„Hacker-Tools“ unter Strafe. Welche Software jedoch unter den Begriff „Hacker-Tools“ fällt, ist im Gesetzestext sehr vage formuliert, sodass erst durch Gerichtsurteile geklärt werden musste, welche Tatbestände tatsächlich unter das Gesetz fallen. Hier ist vor allem die Abweisung einer Verfassungsklage gegen den §220c vor dem Bundesverfassungsgericht zu nennen. Drei Personen hatten geklagt, weil sie – ohne kriminelle Absicht – mit Hacker-Tools umgehen und sich durch den Paragrafen bedroht sahen. Einer der Kläger ist ein Professor, der die Werkzeuge in seinen Lehrveranstaltungen nutzt. Die Verfassungsbeschwerde wurde vom Bundesverfassungsgericht mit Beschluss vom 18. Mai 2009 als unzulässig abgelehnt. Laut Gericht seien „Dual use“-Werkzeuge, die zur Erhöhung der Computersicherheit genutzt werden, aber eben auch für Hackerangriffe taugen, nicht davon erfasst, weil sie einem legalen Zweck dienen. Und wer solche Programme zu erlaubten Tätigkeiten nutze, dem fehle der für eine Strafbarkeit notwendige Vorsatz. Die Richter stellten klar, das selbst ein Ausspähprogramm nur dann strafbar sei, wenn es zum Zweck der Begehung von Straftaten entwickelt oder eingesetzt werde. Dass die Programme geeignet seien, Straftaten zu begehen, genüge nicht. Laut Gesetz sei es auch erlaubt, zu Testzwecken Schadprogramme anzuschaffen. Diese Voraussetzungen lägen bei akademischer Beschäftigung oder bei professionellem Einsatz in der IT-Sicherheit nicht vor.

Allgemein gilt für die IT-Sicherheitsausbildung, dass sie stärker ausgebaut werden muss. IT-Sicherheit sollte im Lehrplan nicht nur optional sein, sondern stärker verankert sein. Es müssen insgesamt mehr Lehrveranstaltungen sowie auf Informationssicherheit spezialisierte Studiengänge angeboten werden. Die offensive Ausbildung bietet eine wertvolle Ergänzung der bisher gängigen Lehrform.

A Bewerbungsbogen zum Kompaktkurs IT-Sicherheit der RWTH Aachen

A.1 Deckblatt

Name, Vorname:
Email-Adresse:
Studiengang:
Semester:
Alter:
Terminpräferenz:
Ich möchte an diesem Kurs teilnehmen weil... :

Allgemeine Hinweise:
Zuerst werden 17 Fragen gestellt, in welchen ihr eine Selbsteinschätzung zu bestimmten Gebieten der IT-Sicherheit wiedergeben sollt. Im Anschluss folgt ein kleiner Wissenstest von 19 Fragen. Bitte beantwortet alle Fragen ehrlich nach Eurem Wissensstand und schlagt keine Antworten nach. Ihr erlangt dadurch keinerlei Vorteile im Kurs. Wählt dann bitte die Antwort „Weiß ich nicht“. Hinweis zum Datenschutz:

Die hier erlangten Daten werden nur für den Kompaktkurs IT-Sicherheit verwendet und elektronisch gespeichert. Personenbezogene Daten werden nach dem Kompaktkurs wieder gelöscht und zu keinem Zeitpunkt an Dritte weitergegeben. Mit dem Ausfüllen dieses Fragebogens erklärst Du Dein Einverständnis zur Speicherung personenbezogener Daten für die Dauer des Kurses. Zur Auswahl einer Antwort bitte ein „x“ in die entsprechende Klammer eintragen. !!! Sofern bei den Fragen nichts anderes angegeben ist bitte nur eine Antwort auswählen

A.2 Awarenessfragebogen

1. Meine Kenntnisse in IT-Sicherheit sind
( ) nicht vorhanden
( ) gering
( ) mittelmäßig
( ) gut
( ) hoch

2. Ich verschlüssele oder signiere meine E-Mails
( ) ja
( ) nein

3. Ich melde mich an meinem Computer als Administrator an
( ) weil es sicherer ist
( ) weil das so eingerichtet war
( ) nein, das tue ich nicht
( ) weiß nicht

4. Dass Online-Transaktionen (Einkäufe, Banking) über eine sichere Verbindung erfolgen ist mir...
( ) sehr wichtig
( ) wichtig
( ) es geht
( ) weniger wichtig
( ) unwichtig

5. Wenn die Möglichkeit zur sicheren Datenübertragung nicht besteht, führe ich den geplanten Vorgang durch
( ) ja
( ) nein
( ) weiß nicht

6. Sicherheitsrelevante Updates meines Betriebssystems führe ich ... durch
( ) automatisch
( ) monatlich
( ) bei Bedarf
( ) selten
( ) gar nicht
( ) Weiß nicht

7. Dass mein PC vor Angriffen geschützt ist, ist mir ...
( ) sehr wichtig
( ) wichtig
( ) es geht
( ) weniger wichtig
( ) unwichtig

8. Ich sperre das Speichern von Cookies auf meinem Rechner
( ) ja
( ) nein

( ) weiß ich nicht

9. Meine Anmeldedaten (Login, Passwort) z.B. für Mailserver speichere ich, um sie später nicht erneut eingeben zu müssen
( ) ja
( ) nein

10. Meine Firewall ist nach meinen Ansprüchen konfiguriert
( ) manuell
( ) automatisch
( ) ich habe keine Firewall installiert

11. Ich erlaube den Zugriff eines Programms auf das Internet, wenn meine Personal Firewall danach fragt
( ) ja, immer
( ) nie
( ) nur wenn ich das Programm halbwegs gut kenne
( ) ich habe keine Firewall

12. Ich schaue nach aktualisierten Versionen meiner installierten Programme, welche nicht automatisch aktualisiert werden oder nicht auf Updates hinweisen
( ) öfter als einmal im Monat
( ) seltener als einmal im Monat
( ) gar nicht

13. Ich erstelle ein Backup meiner wichtigen Daten, d.h. ich sicher diese zusätzlich auf einem externen Datenträger (CD, DVD o.ä.)
( ) öfter als einmal im Monat
( ) seltener als einmal im Monat
( ) nein, ich habe keine wichtigen Daten
( ) nein, ich weiß nicht wie das funktioniert
( ) nein, zu zeitaufwändig

14. Ich verschlüssele meine Festplatte oder bestimmte wichtige Dateien
( ) ja
( ) nein
( ) Ich weiß nicht wie das geht

15. Die größten Bedrohungen/Gefahren für meine Daten sehe ich durch... (alles ankreuzen was zutrifft)
( ) Unbeabsichtigte Fehler
( ) Diebstahl/Einbruch
( ) Gezielte Hackerangriffe
( ) Malware (Viren, Würmer, Trojaner ...)
( ) Technische Defekte an Hardware
( ) Höhere Gewalt (z.B. Feuer, Wasser)

16. Ich benutze regelmäßig folgendes Betriebssystem
( ) Windows
( ) Linux
( ) Beide vorgenannten
( ) Anderes

17. Ich lese Fachzeitungen zum Thema IT-Sicherheit oder informiere mich in Foren, Mailinglisten etc.
( ) Regelmäßig
( ) Manchmal
( ) Nie

A.3 Wissenstest

1. Welche Protokolle übertragen Daten unverschlüsselt? (mehrere richtige Antworten)
( ) FTP
( ) SSL
( ) HTTP
( ) HTTPS
( ) Weiß ich nicht

2. Womit kann der Inhalt eines Verzeichnisse unter Linux angezeigt werden?
( ) grep
( ) ls
( ) mv
( ) Weiß ich nicht

3. Telnet ist besser als SSH wegen ...?
( ) Authentifizierung
( ) Verschlüsselung
( ) Aussage stimmt nicht
( ) Weiß ich nicht

4. Was sind Maßnahmen gegen Passwort Sniffing?
( ) Passwörter dürfen nicht unverschlüsselt/unkodiert per E-Mail gesendet werden.
( ) Passwörter dürfen nicht unverschlüsselt/unkodiert auf irgendwelchen elektronischen Medien gespeichert werden
( ) Passworte dürfen nur verschlüsselt gespeichert werden
( ) Alle Antworten 1 bis 3
( ) Weiß ich nicht

5. Welche der genannten Technologien bietet die potentiell größte Sicherheit für WLAN?
( ) WEP
( ) WPA

( ) WDE
( ) WZA
( ) WEP2
( ) Weiß ich nicht

6. Emails können verschlüsselt/signiert werden mit (mehrere richtige Antworten)
( ) SSL
( ) S/MIME
( ) OpenPGP
( ) SSH
( ) Weiß ich nicht

7. Was ist ein Verschlüsselungsalgorithmus für Festplatten?
( ) Blowfish
( ) Tripwire
( ) SSH
( ) SSL
( ) Weiß nicht

8. Welche Aussage ist wahr bezüglich Netzwerk-Sniffen?
( ) Sniffer erlauben einem Angreifer, Daten die durch ein Netzwerk fließen, zu beobachten
( ) Sniffer ändern die Quell-Adresse eines Computers, um schwache Authentifikationsmethoden auszunutzen.
( ) Sniffer übernehmen Netzwerk-Verbindungen
( ) Sniffer senden sich überlappende IP Fragmente an ein System.
( ) Weiß ich nicht

9. Wie wird ein SYN-Flood-Angriff durchgeführt?
( ) Überhöhte Anzahl Broadcast-Pakete generieren
( ) Hohe Anzahl halb-offener Verbindungen erstellen.
( ) Einfügen periodisch wiederholender Internet Relay Chat (IRC) Nachrichten.

( ) Eine große Anzahl von Internet Control Message Protocol (ICMP)-Nachrichten generieren.
( ) Weiß ich nicht

10. Was wissen sie über Buffer Overflows?
( ) Buffer-Overflows passieren nur auf dem Stack
( ) Geschehen dadurch, dass Eingabe-Daten zum Eingabezeitpunkt nicht auf ihre Länge überprüft werden
( ) Unix-Systeme sind unempfindlich gegen Buffer Overflows
( ) Treten durch unzureichend großen System-Speicher auf
( ) Weiß ich nicht

11. Wofür steht die Abkürzung CIA in der IT-Sicherheit?
( ) Confidentiality Integrity Availability
( ) Conflicts Influences Attack
( ) Control Information Awareness
( ) Control Identification Authentification
( ) Weiß ich nicht

12. Was sind Eigenschaften von Trojanischen Pferden? (mehrere richtige Antworten)
( ) müssen vom Opfer ausgeführt werden
( ) können sich nicht selbständig verbreiten
( ) kopieren sich selbst in nicht infizierte Dateien
( ) reproduzieren sich selbst
( ) müssen nicht vom Benutzer gestartet werden
( ) können mittels Würmern verbreitet werden
( ) Weiß ich nicht

13. Was sind Eigenschaften von Viren? (mehrere richtige Antworten)
( ) Verbreitung über Austausch von Dateien
( ) kopieren sich selbst in nicht infizierte Dateien
( ) verbreiten sich über Sicherheitslücken

( ) versuchen aktiv in neue Systeme einzudringen
( ) Bei einem Virus handelt es sich um ein eigenständiges Programm
( ) Die Wahrscheinlichkeit einer Virus-Infektion unter Linux ist zur Zeit gering
( ) Weiß ich nicht

14. Was sind Eigenschaften von Würmern? (mehrere richtige Antworten)
( ) reproduzieren sich selbst
( ) verbreiten sich über Sicherheitslücken
( ) kombinieren nützliches Programm mit bösartigem Teil
( ) müssen immer zuerst vom Benutzer aktiviert werden
( ) es gibt auch Handy-Würmer, die sich über Bluetooth verbreiten
( ) Ein Wurm ist kein eigenständiges Programm
( ) Weiß ich nicht

15. Was sind Eigenschaften eines Intrusion Detection Systems (IDS)? (mehrere richtige Antworten)
( ) Bestimmt die Quelle eingehender Pakete
( ) Erkennt und meldet Veränderungen an Dateien
( ) Löst Alarm bei bekannten Eindringlings-Mustern aus
( ) Es gibt nur netzwerk-basierte Intrusion Detection Systeme
( ) Angriffe werden nur erkannt, aber nicht verhindert
( ) IDS können selbst nicht Ziel von Angriffen werden
( ) Weiß ich nicht

16. Welches dieser Programme kann für einen Distributed Denial of Service Angriff auf ein Netzwerk verwendet werden?
( ) Satan
( ) Saint
( ) Trinoo
( ) Nmap
( ) Netcat

( ) Weiß ich nicht

17. Was sind Eigenschaften von Rootkits? (mehrere richtige Antworten)
( ) können von Trojanern installiert werden
( ) verbergen Malware vor Antivirenprogrammen
( ) Rootkits können auch zum Systemschutz verwendet werden
( ) Rootkits gibt es nur für Unix-basierte Betriebssysteme
( ) Rootkits gehören zur Klasse der Viren
( ) Rootkits sind Trojanische Pferde
( ) Weiß ich nicht

18. Welcher dieser Angriffe nutzt Buffer Overflows auf dem Betriebssystem aus?
( ) Spoofing
( ) Brute force
( ) Denial of Service
( ) Sniffing
( ) Man-in-the-middle-Angriff
( ) Weiß ich nicht

19. Was ist eine Race Condition?
( ) Ein Standardverhältnis, dass die gesamte Leistung eines eingebetteten Systems ordnet
( ) Der Wettbewerb zwischen zwei Geräten, die die gleiche Unterbrechungs-Priorität teilen
( ) Schutzbedingung die gelten muss, damit 2 Prozesse nicht auf dieselbe Variable schreiben
( ) Konstellation, in der das Ergebnis einer Operation vom zeitlichen Verhalten bestimmter Einzeloperationen abhängt.
( ) Die optimale Bedingung für einen speziellen Interrupt, um die schnellste Ausführungszeit zu erreichen.
( ) Weiß ich nicht

B Awarenesstest nach dem Kurs

1. Wie beurteilen sie ihre Kenntnisse in IT-Sicherheit?
( ) nicht vorhanden
( ) gering
( ) mittelmäßig
( ) gut
( ) hoch

2. Wenn sie bisher noch keine E-Mails verschlüsselt haben, kommt dies jetzt eher für sie in Betracht?
( ) ja
( ) nein
( ) ich verschlüssele bereits

3. Dass Online-Transaktionen (Einkäufe, Banking) über eine sichere Verbindung erfolgen ist mir...
( ) sehr wichtig
( ) wichtig
( ) es geht
( ) weniger wichtig
( ) unwichtig

4. Sicherheitsrelevante Updates meines Betriebssystems werde ich nun ... durchführen
( ) automatisch
( ) monatlich

( ) bei Bedarf
( ) selten
( ) gar nicht
( ) Weiss nicht

5. Dass mein PC vor Angriffen geschützt ist, ist mir...
( ) sehr wichtig
( ) wichtig
( ) es geht
( ) weniger wichtig
( ) unwichtig

6. Meine Anmeldedaten (Login, Passwort) z.B. für Mailserver speichere ich, um sie später nicht erneut eingeben zu müssen
( ) ja
( ) nein

7. Überprüfen sie jetzt öfter als zuvor, ob es aktualisierte Versionen ihrer installierten Programme gibt, welche nicht automatisch aktualisiert werden oder nicht auf Updates hinweisen
( ) ja
( ) nein
( ) weiss nicht

8. Werden sie jetzt öfters als zuvor Backups ihrer wichtigen Daten durchführen?
( ) ja
( ) nein
( ) weiss nicht

9. Werden sie jetzt ihre Festplatte oder Teile davon verschlüsseln?
( ) ja
( ) nein

( ) ich verschlüssele bereits

10. Hat es sie überrascht, wie leicht es geht, ein System zu kompromittieren ?
( ) ja
( ) nein
( ) weiss nicht

11. Empfinden sie ihr aktuelles System als unsicher?
( ) ja
( ) nein
( ) weiss nicht

12. Werden sie Maßnahmen zur Beseitigung von Schwachstellen in ihrem System ergreifen?
( ) ja
( ) nein
( ) mein System hat keine Schwachstellen

13. Die größten Bedrohungen/Gefahren für meine Daten sehe ich durch... (alles ankreuzen was zutrifft)
( ) Unbeabsichtigte Fehler
( ) Diebstahl/Einbruch
( ) Gezielte Hackerangriffe
( ) Malware (Viren, Würmer, Trojaner ...)
( ) Technische Defekte an Hardware
( ) Höhere Gewalt (z.B. Feuer, Wasser)

Platz für weitere Kommentare:

C Aufgabenstellung des Abschlusstests

Name:
IP-Adresse des Rechners:

Sie haben in Ihrer Firma die Administration eines Rechners ubernommen. Auf dem Rechner ist keine grafische Oberflache vorhanden, Sie mussen also auf der Konsole arbeiten. Auf einem Linux-System stehen mehrere Konsolen zur Verfugung. Zwischen den Konsolen kann mit der Tastenkombination $< ALT > - < Funktionstaste >$ n (n=16) umgeschaltet werden.

Von dem Rechner durfen nur die folgenden Dienste angeboten werden: **SSH- und Telnet-Zugang.**
Die Passwortrichtlinie der Firma besagt, dass Passworter mindestens 7 Zeichen haben mussen und aus einer Mischung von Groß- und Kleinbuchstaben mit mindestens einem Sonderzeichen aufgebaut sein mussen. Es gibt keine Vorkehrungen, die die Einhaltung der Richtlinie erzwingen. Es liegt in Ihrer Zustandigkeit als Adminstrator, die Einhaltung der Richtlinie zu uberprufen.

Der vorherige Administrator des Systems hat eine Firewall (iptables) mit einfachen Regeln konfiguriert. Diese soll die folgenden Richtlinien erfüllen:

1. Pakete werden grundsatzlich verworfen.
2. Verbindungen zu dem auf dem Rechner laufenden SSH- oder Telnet-Dienst sind erlaubt.
3. Dem Rechner muss Namensauflosung (DNS) moglich sein,

Hinweis: Die Firewall ist nicht aktiv (und soll auch nicht aktiviert werden). Betrachten Sie das Firewall-Skript in /firewall/. Ihre Aufgabe ist es, das System in einen sicheren Zustand zu bringen. Dazu gehort auch, das System auf mogliche Kompromittierungen zu uberprufen.

- Korrigieren Sie sicherheitskritische Zustande.
- Geben Sie an, falls dies nicht moglich ist. Das Installieren neuer Programme oder des Systems ist nicht erforderlich.

Sie haben zur Bearbeitung der Aufgaben 50 Minuten Zeit !

Dokumentieren Sie Ihr Vorgehen auf diesem Aufgabenblatt, um die Losungen und den Weg dorthin nachvollziehbar zu machen. Am Ende finden sie noch weitere Aufgaben, die Sie aber bitte erst nach der Uberprufung des Systems durchfuhren.

Hinweise: Der Log-Dienst Syslog-ng dient den Organisatoren zur Auswertung des Tests. Er zahlt nicht als Kompromittierung und darf nicht beendet oder entfernt werden.

!!!Wichtig: Arbeiten Sie als root und fuhren Sie direkt nach dem Einloggen auf jeder Konsole den Befehl ' script -a /home/itsec/logX' aus, wobei X der Konsolennummer entspricht, auf der sie arbeiten. (Dies protokolliert Ihre Eingaben) !!!

Beispiel:
Aktion: laufende Dienste uberprufen
begonnen (Uhrzeit): 15.06
Problem: ein WWW-Server ist aktiv
Losung: Dienst WWW-Server abschalten, in dem ...
beendet (Uhrzeit): 15.14

Sicherheitskritische Konfigurationen und Hinweise auf Kompromittierungen

Aktion:
Problem:
begonnen (Uhrzeit):
beendet (Uhrzeit):
Lösung:
(Vorlage kopieren fur mehr Einträge)

Sollten Sie die vorgenannten Überprüfungen vorgenommen haben, dann bearbeiten Sie bitte folgende Aufgaben:

a) Sie möchten auf ihrem System einrichten, dass dem Benutzer betty mittels sudo gestattet wird, die Befehle kill und mount auszufuhren. Welche Zeilen mussen sie in /etc/sudoers einfugen?

b) Konfigurieren Sie den SSH-Daemon. Die geforderten Einstellungen sind:
- Es darf kein direktes einloggen als root möglich sein, sondern nur per su
- SSH soll nicht auf allen Schnittstellen lauschen
- Der Benutzer soll dazu gezwungen sein, ein Passwort einzugeben
- Als SSH Protokoll-Version soll die sichere Variante ausgewählt sein

Literaturverzeichnis

[ACM 2008] ACM: *Computer Science Curriculum 2008: An Interim Revision of CS 2001*. Dezember 2008

[Arce u. McGraw 2004] Arce, Iváan ; McGraw, Gary: Why Attacking Systems Is a Good Idea (Guest Editors' Introduction). In: *IEEE Security & Privacy* 2 (2004), Juli/ August, Nr. 4, 17–19. http://doi.ieeecomputersociety.org/10.1109/MSP.2004.46http://csdl.computer.org/comp/mags/sp/2004/04/j4017.pdf. – ISSN 1540–7993

[Arnett u. Schmidt 2005] Arnett, Kirk P. ; Schmidt, Mark B.: Busting the Ghost in the Machine. In: *Communications of the ACM* 48 (2005), August, Nr. 8, S. 92–95

[Aust 2007] Aust, Christian: Sicherheitsfaktor Mitarbeiter: Aufbau eines Personnel Security Lifecycles für Sensibilisierung, Information, Training. In: *Innovationsmotor IT-Sicherheit – Tagungsband zum 10. Deutschen IT-Sicherheitskongress*, 2007, S. 353–364

[Bacher 1994] Bacher, Johann: *Clusteranalyse. Anwendungsorientierte Einführung*. Oldenbourg, 1994

[Backhaus u. a. 2006] Backhaus, Klaus ; Plink, Wulff ; Erichson, Bernd ; Weiber, Rolf: *Multivariate Analysemethoden. Eine anwendungsorietierte Einführung*. Springer, 2006 (Springer-Lehrbuch)

[Baier u. a. 2003] Baier, Harald ; Buchmann, Johannes ; Busch, Christoph: Aus- und Weiterbildung in IT-Sicherheit. In: *IT-Sicherheit im verteilten Chaos. Tagungsband:*

8. Deutscher IT-Sicherheitskongress des BSI. Ingelheim : SecuMedia-Verlag, 2003, 179–190

[Basel II 2004] Basel II: *Internationale Konvergenz der Kapitalmessung und Eigenkapitalanforderungen.* 2004

[van der Beek 2007] Beek, Frank van d.: *Wie lehrt man IT-Sicherheit am Besten? Eine empirische Studie*, RWTH Aachen, Diplomarbeit, 2007

[van der Beek u. Mink 2008] Beek, Frank van d. ; Mink, Martin: Wie lehrt man IT-Sicherheit am Besten? Eine empirische Studie. In: *Proceedings SICHERHEIT 2008*, 2008, S. 499–511

[BlackHat] *Black Hat Briefings, Training and Consulting.* `http://www.blackhat.com`, Abruf: Mai 2009

[Bogolea u. Wijekumar 2004] Bogolea, Bradley ; Wijekumar, Kay: Information Security Curriculum Creation: A Case Study. In: *Proceedings of Conference InfoSecCD*, 2004

[Bolduan 2009] Bolduan, Gordon: *„Die Russen sind wirklich gut“*. `http://www.heise.de/tr/Die-Russen-sind-wirklich-gut--/artikel/137937/0/0`. Version: Mai 2009, Abruf: Nov. 2009

[Bortz u. Döring 2006] Bortz, Jürgen ; Döring, Nicola: *Forschungsmethoden und Evaluation für Human- und Sozialwissenschaftler.* 4. Auflage. Springer, 2006. – ISBN 3–540–33305–3

[Bräuer 2006] Bräuer, Hermann: Schlendrian mit bitteren Folgen. In: *network Computing* (2006), Oktober. `http://www.networkcomputing.de/schlendrian-mit-bitteren-folgen/`, Abruf: Nov. 2009

[Bühner 2004] Bühner, Markus: *Einführung in die Test- und Fragebogenkonstruktion.* Pearson Studium, 2004

[Campbell u. a. 2007] Campbell, Jamonn ; Greenauer, Nathan ; Macaluso, Kristin ; End, Christian: Unrealistic optimism in Internet events. In: *Computers in Human Behavior* 23 (2007), S. 1273–1284

[Carlson 2004] Carlson, David: Teaching computer security. In: *SIGCSE Bulletin* 36 (2004), Juni, Nr. 2, S. 64–67

[CarmentiS] *CarmentiS – Zusammenarbeit für Frühwarnung in Deutschland*. `http://www.carmentis.org/`, Abruf: Nov. 2009

[Centrum für Hochschulentwicklung 2006] Centrum für Hochschulentwicklung: *Ranking Kompakt: Informatik*. Version: 2006. `http://das-ranking.de/che8/CHE?module=Hitliste&do=show_l1&esb=1&hstyp=1`, Abruf: Nov. 2009

[Chkrootkit] *Homepage "Chkrootkit"*. `http://www.chkrootkit.org/`, Abruf: Nov. 2009

[CIPHER] *Homepage "CIPHER CTF"*. `http://www.cipher-ctf.org/`, Abruf: Nov. 2009

[CISSP] *Homepage "CISSP"*. `http://www.cissp.com/`, Abruf: Nov. 2009

[CompTIA] CompTIA: *CompTIA Security+*. `http://www.comptia.org/certifications/listed/security.aspx`, Abruf: Nov. 2009

[Conti 2005] Conti, Gregory: Why computer scientists should attend hacker conferences. In: *Communications of the ACM* 48 (2005), März, Nr. 3, S. 23–24

[Conti 2006] Conti, Gregory: Hacking and Innovation (Guest Editors' Introduction). In: *Communications of the ACM* 49 (2006), Juni, Nr. 6, S. 33–36

[CSI 2008] CSI: *CSI Computer Crime & Security Survey*. 2008

[Defcon] *DEF CON Hacking Event*. `http://www.defcon.org`, Abruf: Nov. 2009

[Dhillon u. Hentea 2005] Dhillon, Harpal ; Hentea, Mariana: Getting a cybersecurity program started on low budget. In: *ACM-SE 43: Proceedings of the 43rd annual*

Southeast regional conference. New York, NY, USA : ACM Press, 2005. – ISBN 1–59593–059–0, S. 294–300

[Digital Evolution] Digital Evolution: *Homepage "Digital Evolution"*. `http://www.dievo.org/`

[Dimler u. a. 2006] Dimler, Simone ; Federrath, Hannes ; Nowey, Thomas ; Plöss, Klaus: Awareness für IT-Sicherheit und Datenschutz in der Hochschulausbildung – Eine empirische Untersuchung. In: Dittmann, Jana (Hrsg.): *Beiträge der 3. Jahrestagung des GI-Fachbereichs Sicherheit*, 2006, S. 18–21

[Dodge u. a. 2003] Dodge, R. ; Ragsdale, D. J. ; Reynolds, C.: Organization and Training of a Cyber Security Team. In: *Proceedings of the 2003 IEEE International Conference on Systems, Man & Cybernetics*, 2003

[Dornseif u. a. 2006] Dornseif, Maximillian ; Freiling, Felix C. ; Holz, Thorsten ; Mink, Martin ; Anderson, Philip ; Irons, Alastair ; Laing, Christopher: A Comparative Study of Teaching Forensics at a University Degree Level. In: *Proceedings of the Conference on IT Incident Management and IT Forensics (IMF)*, 2006

[Dornseif u. a. 2005a] Dornseif, Maximillian ; Gärtner, Felix C. ; Holz, Thorsten ; Mink, Martin: An Offensive Approach to teaching Information Security: "Aachen Summer School Applied IT Security" / RWTH Aachen. Version: Januar 2005. `http://aib.informatik.rwth-aachen.de/2005/2005-02.ps.gz`. 2005 (AIB-2005-02). – Aachener Informatik Berichte

[Dornseif u. a. 2005b] Dornseif, Maximillian ; Gärtner, Felix C. ; Mink, Martin ; Pimenidis, Lexi: Teaching Data Security at University Degree Level. In: *Proceedings of the IFIP TC11 WG11.3 Fourth World Conference on Information Security Education*, Natalia Miloslavskaya and Helen Armstrong, 2005. – ISBN 5–7262–0565–0, 213-222

[EC-Council] EC-Council: *EC-Council*. `http://www.eccouncil.org/`, Abruf: Nov. 2009

[Farmer u. Venema 1993] Farmer, Dan ; Venema, Wietse: *Improving the Security of Your Site by Breaking Into it.* Usenet Posting to comp.security.unix, 3. Dezember 1993

[Freiling 2009] Freiling, Felix C.: Ein Blick auf IT-Sicherheit aus Angreiferperspektive – Vom Wert offensiver Methoden. In: *DuD – Datenschutz und Datensicherheit* (2009), Nr. 4, S. 214–217

[Freiling u. a. 2008] Freiling, Felix C. ; Holz, Thorsten ; Mink, Martin: Reconstructing Peoples Lives: A Case Study in Teaching Forensic Computing. In: *Proceedings of the 4th International Conference on IT Incident Management and IT Forensics (IMF)* (2008), September

[Freiling u. Mink 2005] Freiling, Felix C. ; Mink, Martin: Bericht über den Workshop zur Ausbildung im Bereich IT-Sicherheit – Hochschulausbildung, berufliche Weiterbildung, Zertifizierung von Ausbildungsangeboten / RWTH Aachen. Version: September 2005. `http://aib.informatik.rwth-aachen.de/2005/2005-20.ps.gz`. 2005 (AIB-2005-20). – Aachener Informatik Berichte

[Gesellschaft für Informatik e.V. 2006] Gesellschaft für Informatik e.V.: *IT-Sicherheit in der Ausbildung – Empfehlung zur Berücksichtigung der IT-Sicherheit in der schulischen und akademischen Ausbildung.* `http://www.gi-ev.de/fileadmin/redaktion/empfehlungen/GI-Empfehlung-IT-Sicherheit-in-der-Ausbildung-2006.pdf`. Version: Oktober 2006

[Hack.lu] *Hack.lu Security Conference.* `http://hack.lu/`, Abruf: Nov. 2009

[Hartung u. a. 2005] Hartung, Joachim ; Elpelt, Bärbel ; Klösener, Karl-Heinz: *Statistik – Lehr- und Handbuch der angewandten Statistik.* Oldenbourg, 2005

[Heise online 2004] Heise online: *MyDoom überlastet Suchmaschinen.* `http://www.heise.de/security/news/meldung/49451`. Version: 27. März 2004, Abruf: Nov. 2009

[Heise online 2008] Heise online: *Microsoft und Uni München: Kampftraining gegen Gefahren aus dem Netz.* http://www.heise.de/newsticker/meldung/110480. Version: Juli 2008, Abruf: Nov. 2009

[Heise online 2009] Heise online: *Dritte Welle der DDoS-Angriffe auf südkoreanische und US-Websites.* http://www.heise.de/security/meldung/Dritte-Welle-der-DDoS-Angriffe-auf-suedkoreanische-und-US-Websites-6119.html. Version: 10. Juli 2009, Abruf: Nov. 2009

[Highland 1992] Highland: Perspectives in Information Technology Security. In: *Education and Society - Information Processing*, 1992

[Hobbs 1868] Hobbs, A.C. ; Tomlinson, Charles (Hrsg.): *The construction of locks / compiled from the papers of A.C. Hobbs, of New York ; and edited by Charles Tomlinson ; to which is added a description of J. Beverly Fenby's patent locks, and a note upon iron safes by Robert Mallet.* London : Virtue and Co., 1868

[HTS] *Homepage "Hack This Site".* http://www.hackthissite.org/missions/, Abruf: Nov. 2009

[Hudec 2003] Hudec, Marcus: *Einführung in die Clusteranalyse.* Version: März 2003. http://homepage.univie.ac.at/Marcus.Hudec/Lehre/WS%202006/Methoden%20DA/Clusteranalyse.pdf, Abruf: 14.07.2007. Skriptum zur Vorlesung Multivariate Statistik

[iCTF] *Homepage "International Capture The Flag".* http://ictf.cs.ucsb.edu/, Abruf: Mai 2009

[Institute of Higher Education 2005] Institute of Higher Education: *Academic Ranking of World Universities – 2005.* Version: 2005. http://www.arwu.org/ARWU2005.jsp, Abruf: Nov. 2009

[Intruded] *Intruded.net Wargames.* http://www.intruded.net/wglist.html, Abruf: Nov. 2009

[Jonsson u. Olovsson 1997] Jonsson, Erland ; Olovsson, Tomas: A Quantitative Model of the Security Intrusion Process Based on Attacker Behavior. In: *Transactions on Software Engineering* Bd. 23 IEEE, 1997, S. 235–245

[Jürjens 2005] Jürjens, Jan: *Sicherheit in der Lehre*. `http://www4.in.tum.de/~juerjens/sicherheit-lehre.html`. Version: November 2005, Abruf: Okt. 2009

[Klauer 2005] Klauer, Karl J. ; Rost, Detlef (Hrsg.): *Das Experiment in der pädagogisch-psychologischen Forschung*. Reprint der Originalausgabe von 1973. Waxmann, 2005. – ISBN 3–8309–1505–5

[Knorz 2008] Knorz, Gerhard: Studie zur IT-Sicherheit – Einstellungen und Einschätzungen zukünftiger Entscheider / Hochschule Darmstadt. 2008. – Forschungsbericht

[Kyas u. a Campo 2002] Kyas, Othmar ; a Campo, Markus: *IT-CRACKDOWN*. mitp-Verlag, 2002. – ISBN 3–8266–0848–8

[Liegl u. a. 2009] Liegl, Marion ; Mink, Martin ; Freiling, Felix C.: Datenschutz in digital-forensischen Veranstaltungen. In: *DuD – Datenschutz und Datensicherheit* (2009), April, Nr. 4, S. 222–227

[Lienert u. Raatz 1998] Lienert ; Raatz: *Testaufbau und Testanalyse*. 6. Beltz, 1998

[Liles u. Kamali 2006] Liles, Samuel P. ; Kamali, Reza: An Information Assurance and Security Curriculum Implementation. In: *Issues in Informing Science and Information Technology* 3 (2006)

[Markoff 1995] Markoff, John: Dismissal of Security Expert Adds Fuel to Internet Debate. In: *The New York Times* (1995), 22. März. `http://query.nytimes.com/gst/fullpage.html?res=990CE7D81739F931A15750C0A963958260`

[Mateti 2003] Mateti, Prabhaker: A laboratory-based course on Internet security. In: *SIGCSE '03: Proceedings of the 34th SIGCSE Technical Symposium on Computer*

Science Education. New York, NY, USA : ACM Press, 2003. – ISBN 1–58113–648–X, S. 252–256

[Mendoza u. a. 2000] Mendoza, Jorge L. ; Stafford, K.L. ; Stauffer, J.M.: Large-sample confidence intervals for the validity and reliability coefficients. In: *Psychological Methods* 5 (2000), Nr. 3, S. 356–369

[Mertens 2007] Mertens, Christian: *Wie lehrt man IT-Sicherheit am Besten – Übersicht, Klassifikation und Basismodule*, RWTH Aachen, Diplomarbeit, 2007

[Mink 2007] Mink, Martin: Ist Angriff besser als Verteidigung? Der richtige Weg für IT-Sicherheitsausbildung. In: *Innovationsmotor IT-Sicherheit – Tagungsband zum 10. Deutschen IT-Sicherheitskongress* Bundesamt für Sicherheit in der Informationstechnik, 2007, S. 339–352

[Mink 2008] Mink, Martin: Über den Nutzen offensiver Lehre. In: *digma - Zeitschrift für Datenrecht und Informationssicherheit* 8 (2008), September, Nr. 3

[Mink u. Freiling 2006] Mink, Martin ; Freiling, Felix C.: Is Attack Better Than Defense? Teaching Information Security the Right Way. In: *Proceedings of the 3rd Annual Conference on Information Security Curriculum Development (InfoSecCD 06)*, ACM Press, 2006, S. 44–48

[Mink u. Nowey 2008] Mink, Martin ; Nowey, Thomas: Human Factors. In: Freiling, Felix C. (Hrsg.) ; Eusgeld, Irene (Hrsg.) ; Reussner, Ralf (Hrsg.): *Dependability Metrics*. Springer Verlag, 2008 (Lecture Notes in Computer Science 4909), S. 188–195

[Näf u. Basin 2008] Näf, Michael ; Basin, David: Conflict or Review – Two Approaches to an Information Security Laboratory. In: *Communications of the ACM* 51 (2008), Dezember, Nr. 12, S. 138–142. `http://dx.doi.org/10.1145/1409360.1409386`. – DOI 10.1145/1409360.1409386

[Neumann] Neumann, Peter G.: *The Risks-Forum Digest*. `http://catless.ncl.ac.uk/risks`, Abruf: Mai 2009

[Neumann 2004] Neumann, Peter G.: Inside risks: the big picture. In: *Communications of the ACM* 47 (2004), September, Nr. 9, S. 112

[Neworder] *Links to Hacking Challenges.* `http://neworder.box.sk/link.php?currentgrp=38667`, Abruf: Nov. 2009

[NSA 2009] NSA: *National Centers of Academic Excellence.* `http://www.nsa.gov/ia/academic_outreach/nat_cae/index.shtml`. Version: 2009, Abruf: Nov. 2009

[Offensive Security] Offensive Security: *Training.* `http://www.offensive-security.com/`, Abruf: Nov. 2009

[OWASP] *OWASP WebGoat Project.* `http://www.owasp.org/index.php/Category:OWASP_WebGoat_Project`, Abruf: Nov. 2009

[Pimenidis] Pimenidis, Lexi: *Hacking Contests.* `http://www.pimenidis.org/research/hacking.php`, Abruf: Nov. 2009

[RFC 1392 1993] RFC 1392: *RFC 1392: Internet Users' Glossary.* Version: Januar 1993. `http://tools.ietf.org/html/rfc1392`

[RKHunter] *Homepage "Rootkit Hunter".* `http://www.rootkit.nl/projects/rootkit_hunter.html`

[Rost 2005] Rost, Detlef: *Interpretation und Bewertung pädagogisch-psychologischer Studien.* Beltz, 2005. – ISBN 3–8252–8306–2

[Ruhr-Universität Bochum] Ruhr-Universität Bochum: *Studiengänge IT-Sicherheit.* `http://www.ei.rub.de/studierende/its/`, Abruf: Nov. 2009

[Sandman 1987] Sandman, Peter M.: Risk Communication: Facing Public Outrage. In: *EPA Journal (U.S. Environmental Protection Agency)* (1987), 21–22. `http://www.psandman.com/articles/facing.htm`

[Scheidemann 2007] Scheidemann, Volker: „It won't happen to me!“ - Aspekte der Risikowahrnehmung mit Anwendung auf den Bereich IT-Sicherheit. In: *Innovationsmotor IT-Sicherheit – Tagungsband zum 10. Deutschen IT-Sicherheitskongress*, 2007, S. 321–337

[Schepens u. James 2003] Schepens, W.J. ; James, J.: Architecture of a Cyber Defense Competition. In: *Proceedings of the 2003 IEEE International Conference on Systems, Man & Cybernetics*, 2003

[Schumacher u. a. 2000] Schumacher, Markus ; Moschgath, Marie-Luise ; Roedig, Utz: Angewandte Informationssicherheit: Ein Hacker-Praktikum an Universitäten. In: *Informatik Spektrum* 6 (2000), Juni, Nr. 23

[Sieber 2008] Sieber, Patrick: *Konstruktion eines Fragebogens zum Thema IT-Sicherheit*. 2008. – Studienarbeit – Universität Mannheim

[Skoudis 2002] Skoudis, Ed: *Counter Hack: A Step-by-Step Guide to Computer Attacks and Effective Defenses*. Prentice Hall PTR, 2002

[Slashdot 2009] *Pentagon Seeks a New Generation of Hackers*. Posting auf Slashdot. `http://it.slashdot.org/article.pl?sid=09/05/22/1627228`. Version: 22. Mai 2009

[Slewe u. Hoogenboom 2004] Slewe, Ton ; Hoogenboom, Mark: Who will rob you on the digital highway? In: *Communications of the ACM* 47 (2004), Mai, Nr. 5, S. 56–60

[Spiegel Online 2007] Spiegel Online: *45,7 Millionen Kreditkartennummern gestohlen*. `http://www.spiegel.de/netzwelt/web/0,1518,474626,00.html`. Version: 29. März 2007, Abruf: Nov. 2009

[Spiegel Online 2009] Spiegel Online: *Bundeswehr baut geheime Cyberwar-Truppe auf*. `http://www.spiegel.de/netzwelt/tech/0,1518,606096,00.html`. Version: 7. Febr. 2009, Abruf: Nov. 2009

[SPSS] *Homepage "SPSS analytics software"*. `http://www.spss.com`, Abruf: Nov. 2009

[Starfleet] *Starfleet Academy Hackits*. `http://isatcis.com/`, Abruf: Nov. 2009

[Steinhausen u. Langer 1977] Steinhausen, Detlef ; Langer, Klaus: *Clusteranalysen*. Gruyter, 1977. – ISBN 3110070545

[Taylor u. a. 2006] Taylor, Carol ; Shumba, Rose ; Walden, James: Computer Security Education: Past, Present and Future. In: *Proceedings of the Seventh Workshop on Education in Computer Security (WECS7)*, 2006

[Tenable Network Security] Tenable Network Security: *Nessus Vulnerability Scanner*. `http://www.nessus.org/`, Abruf: Nov. 2009

[TISP] *Homepage "TISP"*. `http://www.tisp.de/`, Abruf: Nov. 2009

[Vigna 2003a] Vigna, Giovanni: Red Team/Blue Team, Capture the Flag, and Treasure Hunt: Teaching Network Security Through Live Exercises. In: Irvine, Cynthia E. (Hrsg.) ; Armstrong, Helen L. (Hrsg.): *World Conference on Information Security Education* Bd. 253, Kluwer, 2003 (IFIP Conference Proceedings). – ISBN 1–4020–7478–6, S. 3–18

[Vigna 2003b] Vigna, Giovanni: Teaching Hands-On Network Security: Testbeds and Live Exercises. In: *Journal of Information Warfare* 3 (2003), Nr. 2, S. 8–25

[Wälchli 2002] Wälchli: *Sicherheitsmetriken - Übersicht, Analysen und Anwendungen*, Universität Zürich, Diplomarbeit, 2002

[White u. Nordstrom 1998] White, Gregory ; Nordstrom, Gregory: Security across the curriculum: Using computer security to teach computer science principles. In: *Proceedings of the 19th International Information Systems Security Conference*, 1998, S. 519–525

[Wilson u. Hash 2003] Wilson, Mark ; Hash, Joan: Building an Information Technology Security Awareness and Training Program / National Institute of Standards and Technology. 2003 (800-50). – NIST Special Publication

[Wulf 2003] Wulf, Tom: Implementing a minimal lab for an undergraduate network security course. In: *Journal of Computing Sciences in Colleges* 19 (2003), Oktober, S. 94–98

[Yngström u. Björck 1999] Yngström, Louise ; Björck, Fredrik: The Value and Assessment of Information Security Education andTraining. In: *First World Conference on Information Security Education*, 1999, S. 271–292

[de Zafra u. a. 1998] Zafra, Dorothea E. ; Pitcher, Sadie I. ; Tressler, John D. ; Ippolito, John B. ; Wilson, Mark (Hrsg.): Information Technology Security Training Requirements: A Role- and Performance-Based Model / National Institute of Standards and Technology. 1998 (800-16). – NIST Special Publication