GRAY HAT HACKING: MORALLY BLACK AND WHITE

by Courtney Falk

Center for Education and Research in
Information Assurance and Security,
Purdue University, West Lafayette, IN 47907-2086

Title: Gray Hat Hacking: Morally Black and White

Author: Courtney Falk

School: Purdue University, CERIAS

E-mail: court@cerias.purdue.edu

Address: Center for Education and Research in
Information Assurance and Security (CERIAS)
Recitation Building
656 Oval Drive
West Lafayette, IN 47907-2086

Phone: (765) 532-2344

Abstract

This paper sets forth to explore the idea of gray hat hacking – computer hackers outside of an organization breaking into that organization's computer systems with the goal of securing it on behalf of the organization. Gray hat hackers pose a danger because of the uninformed opinions they use to justify their actions. This paper shows similar negative judgments of gray hat hacking from a variety of viewpoints by surveying three prominent normative ethical theories. The target audience is security and computing personnel, managers, supervisors, and others working with computers who may have little to no experience with philosophy and ethics.

1. Introduction

There are persons who take it upon themselves to increase security of other parties' computer systems by breaking into the systems without permission and patching existing security holes in those systems. These people believe they are doing a public service, that this makes their actions morally right. This dangerous fallacy is debunked in the term of this paper by making a survey of three major normative ethical theories, asking if each theory finds gray hat hacking to be morally right or wrong.

There needs to be a distinction of how the word "hacker" is used within this paper. The traditional and popular definitions of "hacker" differ greatly. The traditional definition is used to describe any person building computer hardware or software often in a haphazard and intuitive way but the definition from the popular media is that of a person who breaks into computer systems of other people. A "cracker" is synonymous with both the popular definition of "hacker" and our definition of a black hacker. When this paper refers to a hacker it is speaking about a person who uses unorthodox methods for testing and perhaps penetrating computer security measures.

In the world of computer security there are popular distinctions among system users and their impacts. The distinction is often represented in terms of the color of a hypothetical hat worn by a user. Dr. Eugene Spafford of Purdue University is quoted as saying, "Hats are obvious, behavior isn't. And what is white to one person may be gray to another." To clear up any confusion and to narrow the scope of the paper, the color of a hacker's hat is defined as follows:

- White – typically a professional hired by a company with the specific purpose of maintaining or increasing the existing security of the computer system.
- Black – the stereotypical use of the word "hacker;" a person who breaks into other parties' computer systems without permission and alters data.
- Gray – a person who gains access to other parties' computing systems with the goal of patching and increasing the existing security. Gray hat hackers don't adhere strictly to either of the previously described ideologies but instead set forth to secure other persons' systems because the hackers believe they are doing the right thing. This justification sets up a dangerous situation of unapproved system intrusions.

This paper sets up ethical objections to gray hacking. These objections are raised from three differing normative ethical theories. The three theories discussed are distinct and form the basis for the majority of contemporary ethical theories.

The first theory examined is act utilitarianism as discussed by J.S. Mill (2001). Utilitarianism is an example of a consequentialist moral theory, or a theory concerned with the outcome of actions. The second theory is Kant's deontological moral theory, which is concerned with the will behind the action. The third and final theory is an example of virtue ethics originally proposed by the ancient Greek philosopher Aristotle. While none of the three normative ethical theories covered within this paper have similar focuses they all provide the same interpretation that gray hacking is morally wrong.

2. Utilitarianism

2.1 Mill and *Utilitarianism*

The classic definition of utilitarianism is what is currently referred to as "act utilitarianism." Act utilitarianism (referred to hereafter simply as "utilitarianism"), originally proposed by Bentham and later revised by J. S. Mill in his book *Utilitarianism*, declares that people always act with the pursuit of pleasure (or alleviation of pain) foremost in their minds. This idea is known as psychological hedonism and is central to the teleology of utilitarianism. Teleology tries to explain a phenomenon in terms of a purpose. To utilitarians it is pleasure that is the purpose of action.

Mill summarizes the utilitarian ethic with his greatest happiness principle (p. 7). The greatest happiness principle states that one should act in such a way as to maximize the pleasure (or minimize the pain) for the greatest number of people. This principle is the only rule in Mill's utilitarianism.

Utilitarianism as a consequentialist theory concerns itself with outcomes or consequences of actions and no other aspect of the action such as motive. Therefore one can perform actions that utilitarianism would deem ethically right while still having bad motives. An example would be repaying a debt owed to another person out of fear of physical harm. Repaying a debt is the right action to pursue but fear is a poor reason to do so. This section focuses exclusively on the consequences of actions.

2.2 Application of Utilitarianism

White hackers perform their work in order to create a more secure system or ensure the security of the existing one. While it is doubtful that this will create pleasure it will without a doubt alleviate the pain caused by an insecure system crashing. The consequences of increased security measures and reliability affect all members of the organization that rely upon the computer system. The pain alleviated by the white hacker for the entire group makes the actions of the white hacker the morally right actions to perform.

Black hackers offer a utilitarian interpretation of their actions as straightforward as their white brethren. The black hacker acts in such a way as to create thrills or other personal gain (monetary, social standing, etc.) at the expense of other persons' computer systems. When those computer systems are penetrated and data is tampered with it creates problems both for those charged with protecting the system and those who rely upon the system to complete their own tasks. The creation of pleasure for the individual at the cost of pain for a much larger group of people shows black hacking to be a blatant violation of the greatest happiness principle and therefore is unethical in the utilitarian sense.

Gray hacking offers a challenge to interpret in utilitarian fashion. The altruistic motives enumerated by gray hackers have no place in utilitarianism as the focus is entirely on the outcomes or consequences of actions. But at first glance the gray hacker is thought to be morally right in preventing the pain of network outages for the computer systems' users.

The question of gray hacking's ethicality is whether or not there can be a definitive judgment made on the consequences of the gray hacker creating pleasure or alleviating pain for the most number of people. It becomes apparent that the consequences of this situation are indeterminate. Without prior knowledge and experience with the computer system in question, the actions of the hacker may well break the functionality of crucial software and/or hardware, and because motives are irrelevant in utilitarian ethics any destruction of the computing system, regardless of whether not it is done accidentally, is a bad consequence and a morally wrong action.

However, utilitarianism does not ignore duty, a moral obligation. Mill writes, "Duty is a thing which may be *exacted* from a person, as one exacts a debt (p. 49)." In the situation of the gray hacker it's the system administrators who have a duty to providing the security, integrity, and availability of the computing resources under their control. The debt is very literal because the administrators are taking money in return for their duties of providing computer security. Any gray hat hackers who takes it upon themselves to secure a system for another party are preventing system administrators from fulfilling their moral obligations. Furthermore, by not fulfilling their duties the administrators are not only morally negligent but could lose their jobs for not performing the services for which they are paid. It is now apparent that gray hacking is morally wrong because of its circumvention of others' duties.

2.3 Strengths and Weaknesses of Utilitarianism

The strength of utilitarianism is its simplicity. Resting upon only Mill's greatest happiness principle makes it simple for any person to determine the rightness or wrongness of an action. However there are two serious problems with utilitarianism repeated over the years.

The first problem is that act utilitarianism implicitly advocates the suffering of one person so long as the aggregate pleasure/pain sum is still pleasure. Therefore activities such as slavery, torture, or oppression can be good so long as they generate enough pleasure for enough people even though these ideas offend the considered judgments of any morally mature person who is thinking about those judgments rationally.

A second problem raised with act utilitarianism goes to its teleological nature. A teleological theory defines "good" in terms of some other property, which is pleasure in the case of utilitarianism. For this to hold true all people must agree that pleasure is the only good. Robert Nozick (1971) proposes a thought experiment called the Experience Machine to test this idea. The Machine allows people to plug themselves in and experience any pleasure they desire, similar to themes found in the films *The Matrix* and *The Thirteenth Floor*. Nozick's argument is that if any single person chooses a life not using the Experience Machine then the idea of psychological hedonism fails and without psychological hedonism there is no utilitarianism.

3. Kant

3.1 Kant and the *Grounding for the Metaphysics of Morals*

Immanuel Kant is arguably the most influential western philosopher of the last few centuries. His theory of morals as outlined in his *Grounding for the Metaphysics of Morals* focuses on the motives behind the action instead of the consequences as to what makes something moral or not. This is a deontological theory – one that says goodness is inherent in something. In Kant's theory the good is inherent in the will behind the action. To Kant an action can still be good even if it produces bad results so long as the will with which the action is performed is good.

Kantian ethics revolves around maxims, personal rules by which to conduct one's self. Kant's (1993) central idea is that of a categorical imperative; something any person ought always do no matter what (p. 25). The first formulation of the categorical imperative is referred to as universalizability (p. 14): maxims must be tested by universalizing them, asking if it would still be a desired maxim if everyone would behave in the same fashion (p. 30). Lying is the most common example of a failing the universalizability test because one can't say, "Everyone needs to tell the truth except me," because when the maxim is universalized no one is telling the truth and truth has no value.

Kant divides duty into a four square matrix with categories of perfect/imperfect duties (p. 30) and ourselves/others as parties of action (fig. 1). Perfect duties are required and necessary to life. A perfect duty to others may be not to lie. Imperfect duties are things not required for life but are in one's best interest to pursue. An imperfect duty to one's self might be to take the time and effort in developing his or her own talents.

		Duty	
		Perfect	Imperfect
Party	Yourself	Don’t Commit Suicide	Develop Talents
	Others	Repay Debts	Volunteer

Figure 1.

3.2 Application of Kant’s Theory

White hackers have an imperfect duty to maintain and secure the computer system such that other users within their organization can make use of it. Life continues if they do not perform their duty although their jobs may not. Their motives in protecting the system for others mean they are performing the right actions. Even in the unlikely event that they accidentally cause a failure they are still performing the right actions under Kant’s moral theory.

There is a second formulation to Kant’s categorical imperative stating that one should only treat other persons as ends and never as means to some other end (p. 36). The black hacker breaks this formulation because they abuse the work done by the administrators of a computer system, treating those administrators as means to some other end like pleasure, financial gain, or other forms of personal gain.

The gray hacker follows the second formulation of the categorical imperative closely when they act with the motive of helping other persons. The gray hacker is treating the system administrators as the end and securing the system as a means to that end. But a problem arises when trying to universalize the gray hat hacker’s maxim. Would the same person who breaks in and secures other systems feel the same way when their own computer systems are broken into and secured for them? Quickly it becomes all persons breaking into other people’s computer systems without permission. This lack of respect for other peoples’ property and privacy prevents the maxim from being universalized successfully.

3.3 Strengths and Weaknesses of Kant’s Theory

Kant’s ideas enjoy a large following over the course of several centuries. He appeals to ideas held closely by many people: duty, will, and autonomy, et al. Kant also doesn’t attempt to define “good” in terms of some other property, which is one the focal arguments between naturalistic moralists - those who believe good is definable in terms of other, natural properties - and non-naturalist moralists - those who believe good is a property in its own domain separate from nature - in contemporary ethics (Moore 1903).

The problem with Kant’s theory is its rigidity. To Kant you are always wrong when you lie because it is a perfect duty to others. But what if a person was hiding refugees in his or her home and rebel forces came to the door to round up and execute all refugees and their protectors? (This is a paraphrase of the “Anne Frank Example.”) Few people believe that lying is justified in such a situation. This is an example of simplicity not always providing the best or most elegant solution.

4. Virtue Ethics

4.1 Aristotle and *Nicomachean Ethics*

While being the last of the three normative ethical theories examined, Aristotle’s is also the oldest. According to Aristotle (1999), virtues are characteristics. To possess the virtue is not enough because one must exercise the virtue in action to the proper extent and to the proper person at the proper time (p. 50). Aristotle goes so far as to outline all the necessary parts for a truly good life but this paper focuses only on applying Aristotle's ideas of virtue ethics to the situation of gray hat hacking.

Computer proficiency is considered as a virtue. The ability to examine and exploit computer security measures (i.e. cracking) is also a virtue. But to determine whether or not it is good the person(s), time, and extent to which the virtues are exercised must all be examined.

4.2 Application of Aristotle's Theory

The white hacker uses his/her virtue to the aid of his/her organization. The white hacker's organization is the correct party on which to perform the virtue. An employee of the organization also performs the virtue only to the proper extent needed to test the systems' security. In the event that virtue was exercised in the extreme and caused damage to the systems it is the responsibility of the white hat hacker to repair the damage as part of their duties.

The black hacker is the virtue ethics' antithesis of the white hacker. Black hackers may possess the same virtues as a white hat counterpart but does not care on whom they exercise their virtues. Furthermore, black hackers disregard any proper time for exercising their virtues. Finally, the extent of the virtue exercised is to the extreme such as to allow the black hacker to exploit the computer system to his/her own needs. This category of hacker violates all of Aristotle's necessary conditions for exercising virtues in order to perform a morally right action.

At first glance the rightness of the gray hacker's actions may seem debatable. Examine the extent to which gray hackers exercise their virtues; outsiders often gain access to a computer system by exploiting vulnerabilities in running software, potentially crashing that software in the process. If the web server of a company that relies upon online ordering for their business crashes then the virtues necessary to test the security of the system have been used to the extreme.

There are no right times to exercise these virtues, only bad and worse times. Bad times are those when the network load is light and employees who rely on the system are out of the office and worse times are during peak network usage or work hours.

There are no proper persons on which to perform this action. Grateful system administrators are those who aren't doing their jobs properly and are negligent in allowing outside intruders to secure the system for them. A properly concerned system administrator will spend far more time than the usual workload in tracking down and verifying that the vulnerabilities used in entering the system are patched and also that no other Trojans or back-doors have been installed by the gray hacker. Essentially there is no way for an administrator to know whether or not the gray hacker is in fact gray or a black hacker masquerading as a gray hacker. All three conditions of performing a morally right action are violated by the gray hacker, making gray hacking a morally wrong act.

4.3 Strengths and Weaknesses of Aristotle's Theory

Aristotle's virtue ethics is appealing because it allows for the fact that it is not enough to merely possess a particular characteristic but you must also apply that characteristic to the right person, to the right extent, and at the right time (p. 50).

As an example, the knowledge of CPR is a characteristic one might possess but there are certain conditions on its use. To perform CPR on a completely healthy, conscious person could lead to legal charges. If you break someone's ribs while performing CPR you applied the virtue to the extreme. To perform CPR on a person who has been deceased for hours is futile because it's too late to do the victim any good.

A problem of Aristotle's theory is its lack of any fundamental principles like those put forward by contemporary philosopher Alan Donagan. In the neo-Kantian text, "The Theory of Morality," Donagan (1977) declares the fundamental principle of respecting all persons as rational beings (p. 65). This lack of fundamental principles leads us to situations that may be repugnant such as slavery.

Furthermore there may be objections to what Aristotle (1999) considers virtues. Making bridles and other necessary equipment for riding a horse are coupled together under horsemanship (p. 3). In order to make use of the virtue of horsemanship, like Aristotle suggests, there needs to be at war. It is possible for situations such as this to arise where no good comes from exercising virtues at any time, on any person, and to any extent.

5. Conclusion

This paper examined three different normative ethical theories that represent varying approaches to ethics: consequentialist, deontological, and virtue. The strengths and weaknesses of each theory were also presented as a way of showing how no one theory may be completely sound, but by examining all three together in the survey the strengths of one were used to supplant the weaknesses of another.

While the name "gray hat hacking" seems to suggest moral ambiguity, it is in fact just another form of "black hat hacking" in terms ethical evaluations. Each of the theories presented in this paper from Mill, Kant, and Aristotle show that gray hacking is a morally wrong action and as such should be neither condoned by administrators, managers, or other personnel, nor practiced by well-meaning computer professionals.

6. References

Aristotle. 1999. *Nicomachean Ethics*, trans. M Ostwald. Prentice Hall, Upper Saddle River.

Donagan A. 1977. *The Theory of Morality*. University of Chicago Press, Chicago.

Galouye DF, J Rusnak, and R Centeno-Rodriguez. 1999. *The Thirteenth Floor*, dir. J Rusnak, perf. C Bierko, A Mueller-Stahl, G Mol, and V D'Onofrio. Centropolis Film Productions.

Kant I. 1993. *Grounding for the Metaphysics of Morals*, trans. JW Ellington. Hackett Publishing Company, Inc., Indianapolis.

Mill JS. 2001. *Utilitarianism*. Hackett Publishing Company, Inc., Indianapolis.

Moore GE. 1903. "The Subject-Matter of Ethics." Part I in *20th Century Ethical Theory*, eds. SM Cahn and JG Haber, pp. 12-32. Prentice Hall, Upper Saddle River.

Nozick R. 1971. "The Experience Machine." *Moral Philosophy: A Reader*, ed. Louis Pojman, pp. 124-191. Hackett Publishing Company, Inc., Indianapolis.

Wachowski A, and L Wachowski. 1999. *The Matrix*, dir. A Wachowski, L Wachowski, perf. K Reeves, CA Moss, L Fishburne, J Pantoliano, and H Weaving. Warner Bros.