The Handicap Principle, Strategic Information Warfare and the Paradox of Asymmetry

(Extended Abstract)

Zhanshan (Sam) Ma (University of Idaho)   Frederick Sheldon (Oak Ridge National Lab.)   Axel Krings (University of Idaho)

ABSTRACT[1]

The term *asymmetric threat* (or warfare) often refers to tactics utilized by countries, terrorist groups, or individuals to carry out attacks on a superior opponent while trying to avoid direct confrontation. Information warfare is sometimes also referred to as a type of asymmetric warfare perhaps due to its asymmetry in terms of cost and efficacy. Obviously, there are differences and commonalities between two types of asymmetric warfare. One major difference lies in the goal to avoid confrontation and one commonality is the asymmetry. Regardless, the unique properties surrounding asymmetric warfare warrant a strategic-level study. Despite extensive studies conducted in the last decade, a consensus on the strategy a nation state should take to deal with asymmetric threat seems still intriguing. In this article, we try to shed some light on the issue from the *handicap principle* in the context of information warfare. The Handicap principle was first proposed by Zahavi (1975) to explain the honesty or reliability of animal communication signals. He argued that in a signaling system such as one used in mate selection, a superior male is able to signal with a highly developed "*handicap*" to demonstrate its quality, and the handicap serves "*as a kind of* (*quality*) *test imposed on the individual*" (Zahavi 1975, Searcy and Nowicki 2005). The underlying thread that inspires us for the attempt to establish a connection between the two apparently unrelated areas is the observation that competition, communication and cooperation (3C), which are three fundamental processes in nature and against which natural selection optimize living things, may also make sense in human society. Furthermore, any communication network, whether it is biological network (such as animal communication network) or computer network (such as the Internet) must be reasonably reliable (honest in the case of animal signaling) to fulfill its mission for transmitting and receiving messages. The strategic goal of information warfare is then to destroy or defend the reliability (honesty) of communication networks. The handicap principle that governs the reliability (honesty) of animal communication networks can be considered as the nature's version of "information warfare" strategy because it is a product of natural selection. What we are particularly interested in is to transfer the game theory models (e.g., Sir Philip Sydney game) for the handicap principle to the study of information warfare. In a broad perspective, we realize that the handicap principle may actually contradict the principle of asymmetry in asymmetric warfare. Anyway, not every species of animals has evolved expensive signaling "equipments" like male peacocks (whose exaggerated train is an example of handicap). Furthermore, the handicap principle is not only about communication, it also embodies the spirits of cooperation and competition. In human societies, communication modulates cooperation and competition; the same is true in animal communication networks. Therefore, to evolve or maintain a sustainable communication network, the proper strategy should be to balance (modulate) the cooperation and competition with communication tools (information warfare tools), which is perhaps in contradiction with the asymmetric strategy. There might be a paradox in the strategy of asymmetric warfare, and whether or not information warfare can be used as an asymmetric tool is still an open question.

Keywords: Handicap Principle, Sir Philip Sydney (SPS) Game, Strategic Information Warfare, Asymmetric Warfare, Evolutionary Game Theory.

1. INTRODUCTION

The Handicap principle was first proposed by Zahavi (1975) to explain the honesty or reliability of animal communication signals. He argued that in a signaling system such as one used in mate selection, a superior male is able to signal with a highly developed "*handicap*" to demonstrate its quality and the handicap serves "*as a kind of (quality) test imposed on the individual*" (Zahavi 1975, Searcy and Nowicki 2005). Two examples of the handicap Zahavi cited are the exaggerated train of the peacock and singing in exposed positions by warblers; both are the demonstrations of good quality but they also expose a male to predators. Only males with the superior quality can afford the risk. The handicap principle was initially rejected by scientific community until the early 1990s when evolutionary game modeling showed its validity. Zahavi (1997) later expanded the handicap principle with a monograph titled "*The Handicap Principle: A Missing Piece of Darwin's Puzzle*." The acceptance of the handicap principle changed the landscape of the study of animal communication, and its influence has been recognized gradually outside behavior biology. To some extent, it addresses a question of fundamental importance in both science and humanity—whether or not animal signals are honest or deceitful (i.e., communication reliability)—since it may shed light on the study of trust and altruism in human beings (Searcy and Nowicki 2005).

Computer networks, perhaps any communication networks, are essentially the infrastructures of information, and the goal of network security is to ensure secure and reliable information flow. Uncertainty, vulnerability, and deception are three fundamental properties of communication networks. Information warfare is essentially about the manipulation of these three properties. Vulnerability may be exploited. Uncertainty raises the entropy and lowers the reliability of communication networks. Deception (such as information hiding) can create uncertainty and hide vulnerability.

The objective of this article is to seek inspirations from the research on the principles and mechanism of animal communications (specifically the handicap principle) to the strategy of information warfare. Conceptually, in animal communication networks, communication modulates cooperation and competition between animals; communication also exploits vulnerability, creates uncertainty and utilizes deception strategically and tactically. Mathematically, these properties and processes can be captured with evolutionary game theory models such as Sir Philip Sydney (SPS) game. In fact, game theoretic modeling helped to establish the validity of the handicap principle in biology. If we recognize the fact that uncertainty, vulnerability, and deception are also the fundamental properties of information networks, the evolutionary game theory models devised for studying handicap principle may be applied to the study of strategic information warfare. Nonetheless, due to page limitation, we do not present any mathematical modeling here. One additional aspect we try to explore in this article is whether or not the notion of asymmetric warfare strategy makes sense in the context of strategic information warfare.

2. THE HANDICAP PRINCIPLE

The traditional ethological view of animal communication is that the relationship between a signaler and a receiver is an honest and cooperative information exchange. Dawkin & Krebs (1978) characterized the traditional ethological view as "*it is to the advantage of both parties that signals should be efficient, unambiguous, and informative*." They replaced the ethological view with *non-cooperative* or *manipulative* relationship, i.e., the signaler sends signals to manipulate the receiver's behavior to the signaler's advantage. They stated "*if information is shared at all, it is likely to be false information, but it is probably better to abandon the concept of information altogether*" (Dawkin & Krebs 1978; Searcy & Nowicki 2005).

The partial answer for explaining the dilemma faced by the manipulative theory was already proposed by Zahavi (1975) with the so-termed "*Handicap Principle*." prior to Dawkin & Krebs (1978). Zahavi (1975) argued that in a signaling system such as the one used in mate selection, a superior male is able to signal with a highly developed "*handicap*" to demonstrate its quality and the handicap serves "*as a kind of (quality) test imposed on the individual*" (Zahavi 1975, Searcy and Nowicki 2005). Two examples of the handicap Zahavi cited are the exaggerated train of the peacock and singing in exposed positions by warblers; both are the demonstrations of good quality but they also expose a male to predators. Only males with the superior quality can afford the risk.

The handicap principle did not receive favorable support initially. It was first formulated as a population genetic model, but this model failed to show the evolution of handicaps because the model assumes that only males of high viability can survive with the handicap. This type of handicap is termed "*pure handicap*" or "*Zahavi handicap*" (Searcy and Nowicki 2005). Later two revised versions of handicaps were proposed: *conditional handicap* and *revealing handicap*. The conditional handicap assumes that handicaps are expressed only if a male has *both* handicap and high viability genes. That is, the male is sufficiently viable to endure the additional costs from expressing the handicap in its phenotype.

What leads to the wide acceptance of the handicap principle is the evolutionary game theory (EGT) modeling by Grafen (1990a, 1990b) and Maynard-Smith (1991). The focus of evolutionary game theory (EGT) modeling is often the derivation the so-termed evolutionary stable strategy (ESS). The role of ESS in evolutionary game theory is similar to that of the Nash equilibrium in classical games. ESS is *unbeatable* or *impregnable* in the sense that mutants or dissidents in a population cannot "*invade*" the population under natural selection, in terms of the reduction of fitness. The difference between the Nash equilibrium and ESS is that ESS requires both *equilibrium* and *convergence-stability* conditions (Vincent and Brown 2005). The equilibrium describes the resistance to invasion by a mutant strategy and equilibrium is a necessary condition for ESS. The convergence stability, which is the sufficient condition for ESS, implies that a population will evolve to the ESS when it is near the ESS (Vincent and Brown 2005).

Grafen (1990a) demonstrated that evolutionary equilibrium is reachable when the signaling level is a strictly increasing function of male quality (signal is reliable), and when females prefer males with high levels of signaling (females are responding to the signal). Grafen (1990b) further showed that evolutionary stable strategies are reachable when (*i*) signaling is expensive or signaler fitness declines with the increase of signaling level; (*ii*) receiver's judgment of the signaler's quality increases with the increase of signaler's signaling level, (*iii*) the signaler benefits from the high scores judged by the receiver, and (*iv*) the ratio of marginal cost of signaling to the marginal benefits of a higher level judgment is a decreasing function of signaler quality. Grafen's (1990a, 1990b) game theoretic modeling was still very complicated. Maynard-Smith (1991) later published a two-page short paper that introduced the much simpler Sir Philip Sydney (SPS) game model, but still demonstrated the validity of the handicap principle. A general conclusion from those game-theoretic studies supports the notion that animal signaling can be reliable as long as it is *costly* in an *appropriate* way (e.g., *conditional* or *revealing handicap*).

In the evolutionary game modeling of the handicap principle, the signal cost is the central variable. All costs boil down to *fitness cost* because fitness is the "*hard currency*" to evaluate portfolio in the context of natural selection and evolution of animal communication. According to Searcy and Nowicki (2005), the cost can first be divided into two categories: *receiver-dependent* and *receiver-independent* costs. The former can further be categorized as *vulnerability costs* (e.g., exposed to predation during the signaling), and *receiver-retaliation rules* (e.g., provoke opponent to attack in an aggression conflict). The receiver-independent costs consist of three parts: *production costs* (paid at the time when the signal is displayed to the receiver, e.g., fanning of the train by a peacock), *development costs* (accrued at the time when the signal is developed, e.g., growing elongated feathers to build train by peacock), and *maintenance costs* (need to bear once the signal is developed, regardless of the use, e.g., the reduced flight performance due to the exaggerated train of the peacock). There is yet another classification: *efficacy costs* and *strategic costs* (Maynard-Smith and Harper 2003). Efficacy costs are the costs that produce *good enough* signals but such signals may be undetected by the receiver due to the signal attenuation or background noise. The strategic costs produce signals that cost more than the efficacy costs, but such signals are ensured to be detected by the receiver (Searcy and Nowicki 2005). When costs are identified, the ESS can be computed by balancing the costs and benefits. In the ESS, the deception may exist, not even necessarily in minority. Deception can be in majority as long as the stake (benefits of responding to an honest signal) is sufficiently high enough and the cost of committing deception is sufficient low (Searcy and Nowicki 2005).

The principle of handicap—signals can be reliable as long as they are costly in a proper manner—not only led to the recognition that animal communication can be generally reliable, but also shifted the research focus to issues such as how the signaling could ever be deceptive, and the consequence of cheating (Grafen 1990b, Searcy and Nowicki 2005). Grafen (1990b) analyzed the scenarios in which cheating is tolerable. In the last two decades, several alternatives or extensions to the handicap principle have been proposed (Ma et al. 2009a), but in our opinion, the basic notion of the handicap principle—the balance between cost and signal efficacy—still holds. In our opinion, those alternative hypotheses to the principle can offer meaningful inspiration for studying strategies for information warfare. However, any productive discussion on their implications should ultimately be formulated as mathematical models to be applicable to information warfare. It is an advantage in the case of the handicap principle and information warfare, that evolutionary game models are utilized in both the application domains (Ma et al. 2009b, c). Due to page space limitation, we skip any modeling analysis and only present a few general points in the next section.

3. INFORMATION WARFARE AND THE PARADOX OF ASYMMETRICITY

We suggest searching for inspiration from the handicap principle that is considered to govern strategies of animal communications. The study of animal communication from an information warfare perspective is not new and can be traced back to at least to the now classical work of Trivers (1974) who formulated the parent-offspring conflicts as an information manipulation problem. The investment of parents in offspring care may reduce their own future fitness, and individual offspring may try to maximally taking resources from parents. Offspring is physically weak in taking resources but they can resolve to *information manipulation* (e.g., exaggerated advertisement for their needs). Natural selection must strike a balance to resolve the conflicts. Trivers's (1971, 1974) work, together with several other classical studies including Hamilton (1963, 1964), Axelrod & Hamilton (1981), spawned extensive cross-disciplinary studies characterized by the utilizing the Prisoner's Dilemma (PD) game for the study of cooperation (altruism). Today, there is hardly any field to which game theory is suitable but has not been approached to with the PD game model. In particular, there are extensive theoretical modeling and experimental observations of parent-offspring conflicts in the literature of animal behavior (e.g., Triver 1985, Godfray & Johnstone 2000, Parker et al. 2002, Kilner and Hinde 2008).

In the animal world, information warfare may occur whenever one party has, or contend to have, access to information that their opponents are keen to know (Krebs and Dawkins 1984, Kilner and Hinde 2008). In particular, the asymmetry in physical prowess may motivate the physically weak but knowledgeable rivals to exploit information and get the upper hand (Kilner and Hinde 2008). Therefore, information warfare exists in both nature and human societies. In fact, we realize that the studies of information warfare in both animal behavior biology and information security address fundamentally very similar or even the same questions. But it is also important to understand the distinction of the two fields. One open question that may set the two fields apart is the nature of animal cognition. The handicap principle and its extensions answer the question of whether human beings stand out as a separate group in terms of the truthfulness of their communication. But it does not answer the question of whether or not animals have sufficiently powerful cognitive capability to support the mental states (such as intention and belief). On the other hand, the handicap principle does answer the question that, if the communication is generally honest, what are the mechanisms that enforce the honesty? The answer is the *proper signal cost*. This answer seems to suggest that information warfare in the animal kingdom is also an *economic issue*. For example, as mentioned previously, according to the handicap principle, there are *efficacy costs* vs. *strategic costs* for a communication signal. When the principle is applied to information infrastructure, efficacy and strategic costs may be mapped to reliability and survivability costs, respectively because strategic costs produce signals that cost more than efficacy costs, but such signals are ensured to be detected by the receiver.

The handicap principle also may shed some light on the issue of asymmetry in information warfare. The term *asymmetry* is overloaded even in this article. *Asymmetry* in *physical* *prowess* may motivate information warfare to counterbalance the asymmetry. *Asymmetric warfare* or threats (actions) may inflict disproportionally powerful destructive force to the rivals but may also lead to catastrophic retaliations from the rivals. Research in animal behavior shows that animals rarely use deadly force in their conflict resolutions; instead, they more often use aggression display to establish and maintain the dominance. Asymmetry in *information intelligence* can certainly change the outcomes of information warfare and is often the pursuit goal of information warfare in human societies. But the issue is far from clear in the animal world. For example, *swarm intelligence* of social insects due to simple self-organization mechanism is obviously very effective in terms of the adaptation to environment. A related issue is that swarm *intelligence* is really swarm '*dumbness*' if the term *intelligence* is defined anthropocentrically or brain chauvinistically. It appears that there is always a counterbalance against the asymmetry to evolve (in nature) or maintain (human societies) communication systems. We term this paradox of asymmetry.

The asymmetry paradox and the cost constraints revealed by the handicap principle should not be surprising. It indicates that information warfare is essentially an economic, political, or military matter. In this sense, information warfare has little difference from other kinds of warfare. The asymmetry paradox simply indicates the role of communication in modulating the balance between competition and cooperation, which may mirror the role of diplomacy in mediating the balance between war and peace.

What we discuss in this article may also make sense to general network security, in particular, web security. The web is increasingly mirroring the human society. When some players on the Internet possess asymmetrical advantages, it is difficult, if not impossible, to prevent them from exploiting the power of information. It is not what they declare not to be (evil) that matters. The society may have not recognized the severe consequence of information monopoly to privacy, which perhaps is a much more real threat than the media-popularized information warfare. From a slightly different perspective, invasion to privacy is a kind of asymmetric information warfare against often vulnerable private citizens. It is interesting to observe whether or not one of the lemmas of the handicap principle—economics of the signaling determines its honesty—will resolve the privacy threats from the information monopoly (asymmetry).

Finally, from a broad scientific perspective beyond information warfare, the study of cooperation (altruism) with the Prisoner's Dilemma (PD) game has generated enormous influences across many scientific fields. We argue that the study of communication and handicap principle may spawn important cross-disciplinary ramifications similar to or beyond those generated by the PD game research. For example, the handicap principle, as a principle for *trust* establishment and checking in computer security (particularly semantic web) was suggested in Ma et al. (2009a). Some researchers (e.g., Noe et al. 2001) consider animal societies also run "*market economy*." The handicap principle may just be one of the principles that maintain the order of the *market* in animal societies. It might be interesting to look into the similarity and difference between that market and the *e-market* (e-*commerce*) in features such as the highly distributed nature and the lack of any centralized administrative authority in both markets.

4. REFERENCES

[1] Dawkin, R. and J. R. Krebs. 1978. Animal Signals: information or manipulation? pp.282-309. in "*Behavioral Ecology*", eds. by J. R. Krebs and N. B. Davies. Blackwell, Oxford.

[2] Godfray, H. C. J. and Johnstone, R. A. 2000. Begging and Bleating: The evolution of parent-offspring signaling. Philos. Trans. R. Soc. Lond. B 355, 1581-1591.

[3] Grafen, A. 1990a. Sexual selection unhandicapped by the Fisher process. *J. Theor. Biol.* 144:473-516.

[4] Grafen, A. 1990b. Biological signals as handicaps*. J. Theor. Biol*. 144:517-546.

[5] Hamilton, W. D. 1963. The Evolution of Altruistic Behavior. *American Naturalist*, 97:354.

[6] Hamilton, W. D. 1964. The genetic evolution of social behavior. (I) & (II). *J. Theoretical Biology*. 7:1-52.

[7] Kilner, R. M. & C. A. Hinde. 2008. Information warfare and parent-offspring conflict. *Advances in the Study of Animal Behavior*. pp283-335.

[8] Krebs, J. R. and R. Dawkins. 1984. Animal signals: mind-reading and manipulation. In "Behavioral Ecology". Blackwell.

[9] Ma, Z. S., A. W. Krings, C. C. Hung. 2009a. 'Handicap' Principle for Computer Security, the Semantic Web and Social Networking. *Lecture Notes in Computer Science*, LNCS 5854, pp. 458–468.

[10] Ma, Z. S. 2009b. A New Extended Evolutionary Game Theory Approach to Strategic Information Warfare. *J. Information Warfare* 8(2):25-43. 2009.

[11] Ma, Z. S., A. W. Krings, F. T. Sheldon. 2009c. An outline of the three-layer survivability analysis architecture for modeling strategic information warfare. *Fifth ACM Cyber Security and Information Intelligence Research Workshop*, Oak Ridge National Laboratory. April 13th-15th, 2009.

[12] Maynard Smith, J. 1991. Honest signaling: Sir Philip Sydney game. *Animal Behavior*, 42:1034-1035.

[13] Maynard Smith, J. and D. Harper. 2003. *Animal Signals*. Oxford University Press. 164pp.

[14] Noe, R., J. Hooff, and P. Hammerstein. 2001. Economics in Nature. Cambridge University Press.

[15] Parker, G. A., N. Royce, I. Hartley. 2002. Intrafamilial conflict and parent investment: a synthesis. Philos. Trans. Soc. Lond. B 357:295-307.

[16] Searcy, W. A. and S. Nowicki. 2005. The Evolution of Animal Communications. Princeton University Press. 270pp.

[17] Trivers, R. L. 1971. The evolution of reciprocal altruism. Quarterly Review of Biology, 46:189-226.

[18] Triver, R. L. 1974. Parent-offspring conflict. American Zoologist. 14:249-264.

[19] Triver, R. L. 1985. Social Evolution. Menlo Park, CA. Benjamin Cummings.

[20] Vincent, T. L. and J. L. Brown. 2005. Evolutionary Game Theory, Natural Selection and Darwinian Dynamics. Cambridge University Press. 382pp.

[21] Zahavi, A. 1975. Mate selection–a selection for a handicap. *J. Theor. Biol*. 53:205-214.

[22] Zahavi, A. & A. Zahavi. 1997. *The Handicap Principle: A Missing Piece of Darwin's Puzzle*. Oxford University Press.