Countering Hidden-Action Attacks
on Networked Systems

Tyler Moore

University of Cambridge, Computer Laboratory
15 JJ Thomson Avenue, Cambridge CB3 0FD, United Kingdom
Tyler.Moore@cl.cam.ac.uk

Abstract

We define an economic category of hidden-action attacks: actions made attractive by a lack of observation. We then consider its implications for computer systems. Rather than structure contracts to compensate for incentive problems, we rely on insights from social capital theory to design network topologies and interactions that undermine the potential for hidden-action attacks.

Keywords: computer security, information security, decentralised networks, economics, social capital, asymmetric information, moral hazard

1 Introduction

Recently, cross-disciplinary efforts involving economics and computer security have proliferated [3, 7, 22, 6]. Researchers have typically studied attacks targeting computer systems, and then applied economic principles to develop a deeper understanding of attack properties and defence strategies [3]. In this paper, we take the opposite approach: by turning to the literature on asymmetric information and social capital, we develop an *economic* class of attacks and then study its implications for securing computer systems.

In the theory of asymmetric information, a hidden-action problem arises whenever two parties wish to transact, but one party can take actions that impact the transaction but remain unobservable to the other party. The classic example traces back to the insurance industry, where the insured chooses to behave recklessly (which in turn increases the likelihood of filing a claim) because the insurance company cannot observe the behaviour. This situation generalises to a class of *hidden-action attacks*, which are attractive precisely because observation (and therefore punishment) is difficult or unlikely. In this paper, we focus on this category of hidden-action attacks and attempt to illustrate how these attacks are pertinent to computer network security.

1

Computer networks are naturally susceptible to hidden-action attacks. Routers need not reveal a decision to drop selected packets or falsify responses to routing requests. Nodes can redirect network traffic to eavesdrop on conversations. Users in file-sharing systems can easily hide whether they have chosen to share with others, so many choose to "free-ride" rather than cooperate in the system. In each of these examples, an ability to hide behaviour from other network elements emboldens nodes to carry out attacks.

In the asymmetric information literature, hidden-action attacks are obviated by structuring contracts to induce proper behaviour. For example, deductibles help auto insurers overcome hidden-action problems. By introducing a non-zero cost to file a claim, parties obtain the appropriate incentive for taking reasonable steps to avoid negative outcomes. The need for observation is eliminated, though not without cost: everyone has to pay, even when proper precaution is taken. In computer science, distributed algorithmic mechanism design is attempting to create systems that align all the agents' incentives so that it is in everyone's best interest to operate as intended [27, 18, 10]. However, such an approach is not always practical.

We instead turn to another field of economics in considering how best to deal with hidden-action attacks: social capital. Social capital studies how different institutions are relied upon to facilitate credible transactions among members of a society [19, 23]. In particular, communitarian institutions and markets present a crucial dichotomy.

In a communitarian institution, members from within society cooperate to achieve a goal. For instance, Grameen banks have been established in Bangladesh to make small cash loans in poverty-stricken villages [28]. Traditional banks have failed there because they cannot obtain reliable information about credit risks or monitor progress towards loan repayment. In Grameen banks, entrepreneurs across a village get together to apply for loans as a group. If selected, two members get loans, and once they are successful in meeting the repayment schedule, two more members get loans. The process repeats until the group leader receives a loan. Grameen banks work because group members depend on each other for success, which aligns the incentives to select reliable business partners and monitor each other's progress. Similarly, artisans in 18th-century Britain who could not individually afford to buy land instead contributed to a building society. Once enough money was collected for each plot of land, one society member could build. By pooling resources, building society members could build much faster than they could individually. In contrast, individuals today rely on banks backed by financial markets to issue loans based on largely objective information.

While markets produce more efficient transactions, several important trade-offs must be considered when weighing the efficacy of the systems and their supporting enforcement mechanisms. For instance, the locality afforded by communitarian institutions makes the actions of its participants more observable, diminishing the potential for agents to hide their behaviour. We exploit this fact to deal with hidden-action attacks.

In this paper, we aim to make the security engineer aware of this important

class of attacks so that she may determine whether hidden-action is a credible threat for a given system. When this is the case, she must also be able to find remedies which minimise the threat posed by hidden-action attacks. It is hoped that by presenting the trade-offs between different network topologies and configurations, the task can be made clearer.

We first give a brief overview of the relevant theory from social capital. We then develop an economic model of hidden-action attacks and consider potential countermeasures. Next, we analyse sample hidden-action attacks possible in computer networks, applying lessons from social capital in apprising available mitigation strategies. In particular, we find that existing proposals of reputation systems for peer-to-peer networks are inadequate. Furthermore, they will remain so unless network topology and behaviour are transformed to accommodate decentralised self-enforcement mechanisms. We conclude by discussing open questions and future directions for applying social capital theory to network security.

2 Social Capital

To begin the study of social capital, it is useful to consider the different ways in which people make credible promises to each other [9]. If the parties care about one another, then deciding to make an initially costly commitment is easy. It is also reasonable to suggest that some people are predisposed to trust others [15]. Indeed, in many human societies, members reciprocate whenever they are treated with respect. However, these tendencies only take societies so far. Incentive mechanisms are usually required to facilitate trust-building transactions [8]. Mutually suspicious parties can create credible promises to each other, so long as they can depend on an institution where keeping promises is in everyone's self-interest given that everyone else keeps their end of the bargain. In other words, the institution must facilitate promise-keeping as an equilibrium strategy.

Societies have created vastly different institutions to achieve credible transactions. In each, a mechanism for punishment is used to deter misbehaviour. In many industrialised societies, agreements between parties are explicitly translated into a contract whose legitimacy is supported by a separate authority called an *external enforcer*. An external enforcer's legitimacy stems from the need to maintain a solid reputation among society's members. Conducting regular elections and establishing a free press to investigate corruption are ways to ensure that an external enforcer must consider its reputation [21].

In many societies, people cannot rely upon the legitimacy of contracts to transact because no credible legal system exists for enforcement. However, transactions can and do occur. A primary requirement is for group members to face repeated opportunities to transact; then the members of the group could enforce agreements themselves [9]. *Mutual enforcement* mechanisms rely on a credible threat by all community members to punish anyone who breaks the agreement (by refusing future transactions, for example). So long as the wind-

fall from cheating in one round is offset by the expected loss from future rounds, mutual enforcement mechanisms suffice.

We can now define social capital. (Social capital is a lively topic throughout the social sciences, and it is therefore no surprise that myriad definitions abound.) We adopt a minimalist definition espoused by Dasgupta, namely, that social capital is simply a system of interpersonal networks connecting members of a group [9]. Closely related to the notion of social capital is a *resource allocation mechanism* that defines interactions between network members. Dasgupta is keen to separate the resource allocation mechanism from social capital because different mechanisms can be used on the same interpersonal network with varying consequences. Some societies, notably less developed ones, rely on these networks more than others to reach credible agreements. Different mechanisms reflect different ways for trust to operate among members.

2.1   Resource Allocation Mechanisms

We consider two representative resource allocation mechanisms: markets and communitarian institutions. Markets found in industrialised societies operate under the watch of an external enforcer. Interpersonal networks are not required for operation in market institutions; rather, the parties to a transaction may appear anonymous. (Of course, names and reputation can influence a transaction, but we are considering the ideal case.) Anyone can purchase goods provided they have the means for payment; likewise, people are free to choose which goods to produce. Because personality does not matter, markets are capable of achieving allocative efficiency—production moves to the most efficient party.

External enforcement mechanisms naturally support market institutions by reinforcing attributes critical to markets. For instance, just as markets do not differentiate between parties to a transaction, external enforcement mechanisms rely upon a legal structure which treats parties equally. By keeping the enforcement mechanism independent from the parties to a transaction, agents only need faith in the enforcement mechanism, not in one another. Repeated interactions with each other are unnecessary; instead, repeated interactions with the external enforcer facilitate trust. Such a system complements markets well since substantial populations interacting in a common market are unlikely to anticipate future interactions.

Communitarian institutions, on the other hand, directly depend upon interpersonal networks for establishing trust. These institutions are crucial in less developed economies, such as isolated villages in rural India. Parties invest considerable time and effort into existing channels, so they are naturally reticent to endure additional cost by seeking out new transaction partners. One consequence is high value placed on personalities during transactions. Another is potentially inefficient resource allocation. Of course, relying upon connected members for trade is a suboptimal technique when compared to markets. This is not surprising given that the same mechanism meant to establish credibility is used to allocate resources.

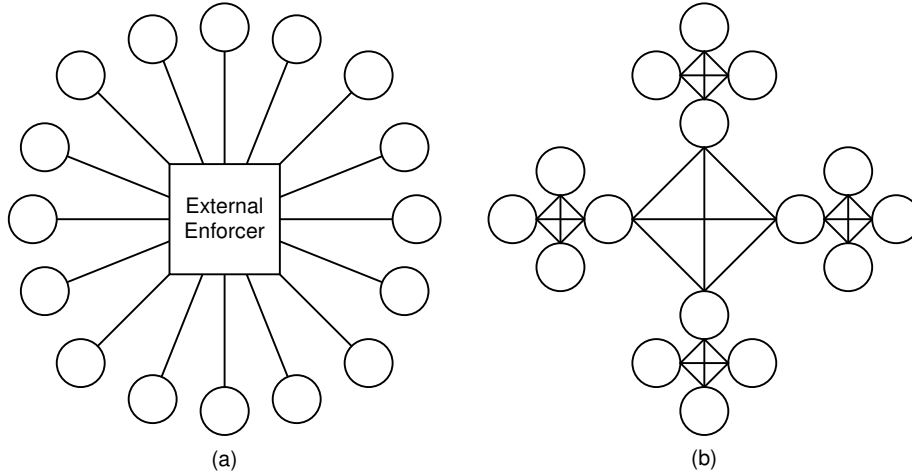Mutual enforcement mechanisms function well under the constraints of a

4

Figure 1: Network topology for (a) market and (b) communitarian resource allocation mechanisms

communitarian institution. A relatively small group size ensures that members are likely to interact repeatedly. Furthermore, members can keep a watchful eye on the behaviour of their transaction partners, ready to punish violations at a moment's notice. Ease of observation provides an effective deterrent to would-be violators and colluders. Note that the threat of punishment extends recursively: refusing to inflict punishment for a violation yields punishment, and so on.

2.2   Observability and Network Topology

Figure 1 highlights the different interpersonal networks relied upon for the representative resource allocation mechanisms. Market institutions are indicated by a star topology graph since all transactions are (implicitly) mediated by an external enforcer (Figure 1(a)). Nodes need not observe each other's actions; instead, relevant information can be noted by the external enforcer. Whenever conflicts arise, the parties must turn to the enforcer for resolution. Therefore, transactions are documented so that the enforcer can verify member actions.

Such an enforcement strategy deters many violations and enables external enforcement techniques to operate effectively on large populations. However, it also underscores a fundamental difference between the capabilities of the enforcement mechanisms and the corresponding resource allocation mechanisms. Since not all activities can be observed, markets are susceptible to abuse whenever the actions of an agent can be kept from its transaction partner and the external enforcer.

Because observation is endogenous to communitarian institutions, they do not suffer the same problem. The graph in Figure 1(b) contains smaller, fully-

connected clusters which are then connected to other clusters. No external enforcement mechanism mediates; rather, each node incorporates the threat of punishment into its equilibrium strategies. Note that the capability for efficient observation is incorporated into the topology.

The edges connecting the nodes together represent distribution of social capital for the two representative societies. In markets, transactions can occur between any pair of nodes, but the trust is established between the nodes and the external enforcer, not the transacting nodes themselves. Repeated interactions with other nodes are unlikely given the population size. With communitarian institutions, fewer channels are established but they are relied upon repeatedly for trade. As a result, nodes are limited by the scope of trading partners, but the locality of connected channels makes mutual enforcement feasible.

3 Hidden-Action Attacks and Countermeasures

A key advantage of communitarian institutions over markets is each agent's ability to observe the behaviour of its neighbours consistently. This distinction inspires a category of attacks. Hidden-action attacks are precisely those actions made feasible by a lack of sufficient observation. In terms of social capital, hidden-action attacks are deviations made attractive when using a market-style resource allocation mechanism but not when operating communitarian ones.

3.1 Hidden-Action defined

We construct a simple model for hidden-action attacks. The aim is to give a precise definition so that we can reason about ways to deal with this attack category. Hidden-action is best understood by comparing the utility-maximising strategies of agents in systems where observation is unlikely to systems where it is likely.

First we shall define the expected utility for an agent engaging in a transaction. She can choose to abide by (A) or break (B) the parameters of the agreement. We consider behaviour in two mechanisms—one where observation is difficult (e.g., market mechanism backed by external enforcement, call it m) and one where observation is easy (e.g., communitarian institution mutually enforced, call it c).

$$u_A = v_A - d_A$$

$$u_B = v_B - d_B - P(\text{detection}|B) * \text{penalty}$$

Here, v represents the value and d the disutility of the chosen action. We assume that it is more costly to cooperate than deviate ($d_A > d_B$) and more valuable individually to deviate ($v_B > v_A$). It is safe to assume that the probability of detecting an attack is higher when actions can be observed ($P_c(\text{detect}) > P_m(\text{detect})$).

In order for hidden-action attacks to be viable, two conditions must be met: (1) the agent must be better off deviating when the likelihood of detection is

low, and (2) the agent must be better off cooperating when the likelihood of observation is high:

$$v_B - d_B - P_m(\text{detect}|B) * \text{penalty}_m > v_A - d_A \qquad (1)$$

$$v_B - d_B - P_c(\text{detect}|B) * \text{penalty}_c < v_A - d_A \qquad (2)$$

These conditions describe a situation where *only* a low likelihood of detection encourages an agent to deviate. When faced with a system where observation is likely, the agent changes its behaviour. Rearranging the inequalities, we arrive at a definition for hidden-action attacks.

Definition 1 *An action B is considered a* hidden-action attack *whenever its benefits and costs to an agent satisfy the following inequalities:*

$$P_m(\text{detect}|B) * \text{penalty}_m < (v_B - d_B) - (v_A - d_A) < P_c(\text{detect}|B) * \text{penalty}_c$$

This definition says that hidden-action attacks may occur whenever the net utility gain from deviating lies between the expected penalty enforced when observation is unlikely and the expected penalty enforced when observation is likely. If the expected gain in attacking does not exceed the expected penalty even when actions can be hidden, then no attack should occur. And if the expected gain in attacking exceeds the expected penalty even when observed, then the attack may be launched regardless. Even if observation diminishes the attacker's net gain, the penalty is not a large enough deterrent; thus the attack does not rely on hidden-action. Such a definition for hidden-action attacks makes sense—these actions are attractive to an attacker only when she knows observation is difficult or unlikely. In the following section we consider ways to make observation easier.

3.2   Countermeasures

We have deliberately avoided adopting an approach traditionally taken by economists in countering the effects of asymmetric information—principal-agent theory. In principal-agent theory, contracts are devised to compensate agents capable of hidden-action. Agents receive payments to align their incentives with those of the principal. In the case considered here, agents could receive payments to make cooperating in transactions preferable to deviating in an unobservable way. In certain circumstances, payment schemes are an effective technique for aligning incentives. However, we reject this approach in our analysis because payment schemes are often burdensome and difficult to implement, especially in computer networks.

Another difficulty in relying upon principal-agent theory as an analytical tool is that it views hidden-action as an unchangeable property of the system. Rather, we hope to uncover effective ways to minimise or even eliminate hidden-action attacks by designing network interactions accordingly. Therefore, designing incentive contracts is better viewed as a last resort for deployed systems where architectural modifications are impossible.

Furthermore, the theory of social capital offers a compelling alternative to incentive contracts: increase observability by transforming the ways agents interact. We consider the properties of resource allocation mechanisms when considering countermeasures.

Social capital tells us that the likelihood of verifying that an attack has occurred in a market allocation mechanism is significantly lower than the likelihood of observing an attack when a communitarian allocation mechanism is used. In fact, a common simplifying assumption is that, for hidden-action attacks, verification by an external enforcer is impossible in pure market mechanisms while observation is perfect when mutually enforced by a watchful community. In reality, most resource allocation mechanisms fall somewhere between the two extremes. Therefore, any strategy to reduce the number of feasible hidden actions should attempt to increase the observability of agent actions. But how can this be achieved? Adhering to the discussion from Section 2, communitarian institutions incorporate observation into the network topology and rely on it to carry out behavioural enforcement.

Constructing a network as a series of connected neighbourhoods rather than a flat, widely interconnected system (see Figure 1(b)) creates the potential for accountability by incorporating locality into the network structure. Nodes belonging to a cluster should depend upon one another by interacting frequently. This increases the level of trust through persistent observability. Repeated interactions are critical for effective observation; otherwise, mutual enforcement is ineffective because the discount rate becomes unworkably large.

Next, shifting enforcement responsibilities to the members themselves ensures that the incentives to keep a watchful eye on transactions are aligned. Shifting punishment responsibility is not enough, though. Each member must be able to count on other members to dedicate resources to watching so that an equilibrium strategy can be reached.

4   Hidden-Action in Computer Networks

Having identified an economic category of hidden-action attacks, we now consider its potential impact on computer networks.

A fundamental challenge to securing computer systems arises whenever systems are connected together. The Internet has created a globally addressable network, interconnecting diverse domains accessible by all. It is no surprise, then, that the capability for hidden-action across the Internet is quite considerable. It is often impossible to trace the source of Internet attacks, so the expected penalty does not serve as an effective deterrent [22]. However, many computer network applications are capable of constraining population size, actions and other properties of interconnected nodes. In such systems, hidden-action attacks may be overcome by incorporating lessons from social capital into system design.

4.1 Hidden-Action Attacks

Many computing systems inadvertently enable hidden-action attacks through design choices. For instance, network topology and shared responsibility enable hidden-action attacks in sensor networks. Many sensor networks designate particular nodes the task of aggregating reported values, e.g., temperature readings, from several sensors. The aggregating node then relays the compiled value to a base station across the network. The aggregating node could easily report a false value to the station. Its actions would be hidden from the other sensors because only an aggregated value is reported. An asymmetry of responsibility combined with a lack of transparency makes this hidden-action attack possible. A simple countermeasure for distributing responsibility is to periodically rotate the task of aggregating across nodes, while having a second node redundantly compute the aggregated value increases observability. These techniques reinforce the requirements for a mutually-enforced, communitarian-style institution.

Another example hidden-action attack is a router selectively forwarding messages [11]. It is the router's job to forward messages to their proper destinations. However, a malicious router may choose to drop certain messages it deems undesirable. A router can hide behind the fact that failures are expected to happen occasionally. The extent to which this attack is hidden depends upon the stability of the routing topology, which impacts the likelihood of repeated interactions. Telecommunications signalling protocols (e.g., SS7) often use statically defined routing tables. Here, repeated interactions occur frequently so any persistent failures will be quickly noted. At the other extreme, ad hoc routing protocols define routing paths that may be used just once. Such a high discount rate for future interactions enables lazy routers to drop traffic irrelevant to itself and allows malicious ones to selectively forward traffic or propagate false information. Backbone routers along the Internet fall somewhere in between by using dynamic routing tables that rarely change significantly.

Redirecting network traffic to eavesdrop is a particularly pernicious hidden-action attack because the victim experiences no change in functionality. The extent to which such attacks are feasible depends upon population size, network topology and distribution of routing responsibility. If a system relies upon a single node to route network traffic, e.g., a media gateway controller that establishes call streams in a Voice over IP network, then increasing transparency of its actions so that network members can easily monitor them can eliminate the potential for hiding its behaviour.

Peer-to-peer (P2P) systems have suffered from widespread "free-riding," where a user draws on a system's resources without contributing back later. We consider this to be a consequence of the great potential for hidden-action in peer-to-peer systems. In the following section, we analyse more closely the implications from social capital for addressing this hidden-action problem.

4.2   Hidden-Action in Peer-to-Peer Systems

Peer-to-peer systems represent a new, decentralised paradigm for computing where nodes have dual roles as clients and servers [2, 26]. Unfortunately, no effective mechanisms have been put in place to ensure that peers balance their responsibilities to both roles. Free-riding in deployed systems is well-documented [1], while academic approaches have failed to find an adequate solution given the constraints of the network. We claim that free-riding remains viable because peers can easily hide their actions from others. Furthermore, susceptibility to such hidden-action stems from the structure of the network topology and node interactions.

It is helpful to consider relevant assumptions for transactions in peer-to-peer systems. First, P2P systems exploit network externalities to the fullest extent by accommodating large member populations with a flat network topology. In fact, joining one creates the potential for collaboration with every other peer in the system. High turnover is also expected; nodes may join and leave the system in a matter of minutes. These properties make the prospects for repeated interactions very discouraging. Inexpensive or even costless identities further exacerbate the problem of unrepeated interactions while also making penalties more difficult to implement. Given a network design with these properties, nodes are certainly predisposed to hidden-action. The attributes given above coincide with many of those of market-based institutions presented in Section 2.1, but without an adequate external enforcement mechanism to deter misbehaviour.

In fact, much of the research on overcoming the freeloading problem has focused on developing a viable *mutual* enforcement mechanism. Mutual enforcement is preferred for its scalability and decentralisation. However, such attempts directly contradict the lessons of social capital research: namely, mutual enforcement mechanisms require (1) repeated interactions, (2) far-sighted nodes and (3) sufficient capability to punish deviation. In their present form, P2P networks meet *none* of these requirements. Instead, many of the efforts to overcome freeloading are actually attempts to add exogenous features in hopes of meeting the requirements for mutual enforcement.

For example, Feldman et al propose maintaining a network-wide shared history of past transactions to simulate repeated interactions and raise the level of observation [12]. Yet any such shared history approach is susceptible to gaming by malicious users when the population is large and dynamic. They also outline a reciprocative strategy for dealing with newcomers that attempts to punish free identities without penalising too severely legitimate users who join the system. The results are mixed at best: while demonstrating marked improvement over P2P systems with no enforcement mechanism, these systems still break down whenever the population grows too large. This negative outcome reinforces results from social capital since existing P2P system attributes contradict the requirements for effective mutual enforcement mechanisms.

So what options remain to overcome free-riding? Is mutual enforcement of P2P systems even a realistic goal? According to social capital theory, not in their present form. However, fundamental changes to network topology may offer a

solution. Currently, when a peer joins the system, he can transact with anyone throughout the network. While beneficial in that it maximises transaction possibilities, a universally accessible network topology makes remote observation difficult and repeated transactions unlikely.

An alternative is to adopt a network topology resembling connected neighbourhoods of nodes. Here, nodes must first transact with other members of the neighbourhood to establish legitimacy. Once trust has been established inside the clustering of nodes, outside transactions can occur through established channels between groups. Such a network topology facilitates self-enforcement by establishing a credible threat of observation to forestall hidden-action. To improve efficiency, network behaviour must be designed so that group members collaborate with each other for most common transactions. One way to achieve this is to structure group membership around particular interests. Specialised groupings naturally facilitate communitarian-style institutions, and they can even improve search and communication efficiency by reducing the need for contact with the rest of the system.

4.3   Lessons for the Security Engineer

What resources are available to a security engineer for dealing with hidden-action attacks? She must first consider whether nodes can easily hide actions in the network. When this is the case, any solution must increase the threat of observation. Changes to network structure and operation, as exemplified in the above sections, are often required.

Interactions must be designed so that the network members concerned in a transaction can monitor its outcome at a reasonable cost. Most monitoring systems are exogenous to the network under observation, that is, monitoring is an additional feature added after the system has already been developed. To effectively overcome hidden action attacks, however, networks must be designed so that monitoring is endogenous to the system. Building locality into the network topology ensures that a credible threat of observation is maintained. Some transactions require collaborating with distant nodes; for those that do not, placing nodes in close proximity can aid in mutual observation. If nodes join a large network with unstable membership, then establishing a low-cost means of monitoring activity is difficult. Reducing uncertainty in anticipated transaction properties and partners helps facilitate endogenous monitoring.

Likewise, no network member should be expected to provide services without also depending upon another member for service. This principle can be incorporated into the responsibilities placed upon nodes. Mutual dependence strengthens the likelihood of repeated interactions, as well as fostering an incentive to protect correct network behaviour.

5 Discussion and Open Questions

Some remaining subtleties and outstanding questions must be considered to yield a deeper understanding into the application of hidden-action to computer networks.

So far, we have focused on just the positive implications of communitarian-style institutions over markets in addressing hidden-action attacks. However, several fundamental limitations deserve mentioning.

First, constructing network topologies as connected neighbourhoods may increase observability and raise the level of repeated interaction, but this comes at a price: inefficient resource allocation. Intuitively, this makes sense. Consider a P2P file sharing system. The libraries of several users in neighbourhood can never be as diverse as the compiled library of all the users of the system. Applications must balance the need for allocative efficiency against the potential harm of hidden-action.

Another negative implication of neighbourhood topologies is the tendency towards risk correlation. Insurance markets reach efficiency by using large populations to create independent risks. However, neighbourhoods of nodes are likely to be susceptible to many of the same risks. But what are the implications for computer networks? It is true that one malicious node is more likely to wreak havoc on its neighbours than the rest of the network. But what if the increase in repeated interactions and threat of monitoring makes deterring many such attacks more feasible? Can we arrive at a network-wide welfare analysis comparing the expected utilities of neighbourhood network topologies versus universal ones?

We have also treated privacy concerns as a non-issue. In many ways, privacy and observability are mutually exclusive concepts. Whenever members value the privacy of a particular action, it can remain hidden. However, such actions are then vulnerable to hidden-action attacks, and so group members must weigh the expected costs of these attacks against the value of privacy for the given action. One reason privacy erosion in computer networks is so alarming is that personal details and actions can be made available across a very large network, often the Internet. Decentralised observation by a limited number of relatively trusted nodes is a potentially much smaller problem for many applications, and it also makes pseudonyms more workable through repeated interactions.

The dichotomy of resource allocation mechanisms we have considered—markets supported by external enforcement versus communitarian institutions mutually enforced—represents two extremes. In reality, many systems often incorporate features from both. Powell and Brantley point out that researchers working for rival biotechnology firms share some information with one another while keeping other information secret [20]. These interactions reflect a complementary role between markets and interpersonal networks [9]. Indeed, the value of collaboration between scientists explains the proliferation of groupings of firms in common geographical locations, such as Silicon Valley in California and the Research Triangle in North Carolina. Similarly, systems designers must strike a balance between the two representative mechanisms to arrive at a level

of observation sufficient to overcome threatening hidden-action attacks while maintaining the greatest allocative efficiency possible.

Much of the recent systems research has focused on decentralised networks (e.g., P2P, sensor networks). Is mutual enforcement the only viable mechanism for deterring misbehaviour in these systems? Can external enforcement be deployed without resorting to centralisation? More research is needed to determine the viability of verifiable, decentralised audit and accounting mechanisms. Furthermore, could a strengthened Internet law ever serve as a viable external enforcement mechanism? Hidden-action problems would not go away, but they might be marginalised.

The assumption of no-cost monitoring in communitarian institutions is compelling when considering people transacting in a small village. But monitoring does bear costs in computer networks, especially when nodes interact frequently. Physical locality is not a compelling argument to support low-cost monitoring in networks. Systems can be physically separated across the globe—what matters is logical locality, i.e. whether the nodes share the same network addressability and can examine each other's network communications. How can we model the true costs of monitoring?

6   Related Work

While not addressing hidden-action attacks directly, many computer security researchers have identified a need to align the incentives of nodes in a network, particularly among nodes in peer-to-peer [12, 25, 13, 17] and ad hoc networks [4, 5]. Proposed remedies fall into two broad categories: payment schemes and reputation systems.

Payment schemes take a fairly direct approach, compensating users for the costs undertaken in performing tasks [2, 5]. Users earn credit for performing costly operations which can later be redeemed in exchange for cooperation from other nodes. Difficulties arise whenever legitimate users are positioned at the edges of networks where few costly actions are required. Furthermore, these systems are quite complex and introduce significant overhead into a network.

In [11], Feldman and Chuang characterise multi-hop routing attacks where routers drop selected packets as a hidden-action problem. They devise contracts consistent with principal-agent theory to overcome the routers' incentive to drop packets. By assuming a router's actions can never be observed, making such contracts becomes the only viable solution. In contrast, we have investigated ways to structure networks so that actions can be made observable.

Reputation systems attempt to keep track of the behaviour of nodes to create a credible threat of punishment for misbehaviour [16, 12, 4]. Unfortunately, as this paper has demonstrated, these systems face an uphill battle given the constraints of the networks they are deployed on—large populations, low likelihood of repeated interactions and decentralised enforcement (see Section 4.2). Furthermore, these systems are susceptible to collusion, and social choice theory casts significant doubts on the prospects of overcoming these problems in their

present design [24]. We instead turn to social capital to identify critical ways to alter network composition and behaviour.

The questions posed here relate to an ongoing debate in social policy—does solidarity or diversity better serve society [14]? Diversity encourages fresh ideas, but solidarity enables stronger bonds between members of a community. In [7], Danezis and Anderson argue that users in peer-to-peer systems are better equipped to defeat censorship by choosing what to share rather than serving a random sampling. We also side with solidarity by carrying this concept a step further: structure network topology and interactions so that like-minded nodes are more likely to transact repeatedly.

7 Conclusions

We have defined an economic category of hidden-action attacks and demonstrated its application to computer systems: from nodes in a sensor network falsifying aggregated readings to users in a peer-to-peer system "free-riding" off unsuspecting contributors. Instead of devising contracts to compensate misaligned incentives, we have adopted results from social capital research and identified ways of designing the network topology and interaction of computer systems to undermine the allure of hidden-action attacks.

Recent research efforts into peer-to-peer systems and sensor networks highlight a trend towards decentralisation and mutual enforcement. We have shown that the environmental assumptions of decentralised networks often contradict the requirements for effective mutual enforcement mechanisms. To accommodate mutual enforcement techniques, we have proposed constructing decentralised networks as connected neighbourhoods, using communitarian institutions as a model.

We hope to have demonstrated that the ways in which human institutions are structured present meaningful lessons for computer systems designers. Much work remains to be done, however, in comprehensively identifying the range of potential hidden-action attacks on computer systems. Perhaps the next challenge is to develop and deploy networked systems that incorporate attributes from communitarian institutions.

References

[1] E. Adar and B. A. Huberman. Free riding on Gnutella. *First Monday*, October 2000.

[2] R. Anderson. The eternity service. In *First International Conference on the Theory and Applications of Cryptology (PRAGOCRYPT '96)*, 1996.

[3] R. Anderson. Why information security is hard—An economic perspective. In *ACSAC '01: Proceedings of the 17th Annual Computer Security Applications Conference*, page 358. IEEE Computer Society, 2001.

[4] S. Buchegger and J.-Y. L. Boudec. Nodes bearing grudges: towards routing security, fairness, and robustness in mobile ad hoc networks. In *Proceedings of the Tenth Euromicro Workshop on Parallel, Distributed and Network-based Processing*, pages 403 – 410, Canary Islands, Spain, January 2002. IEEE Computer Society.

[5] L. Buttyán and J.-P. Hubaux. Enforcing service availability in mobile ad-hoc WANs. In *MobiHoc '00: Proceedings of the 1st ACM international symposium on Mobile ad hoc networking & computing*, pages 87–96. IEEE Press, 2000.

[6] J. Camp and S. Lewis, editors. *Economics of Information Security*. Kluwer, 2004.

[7] G. Danezis and R. Anderson. The economics of censorship resistance. In *Proceedings of the Third Annual Workshop on the Economics of Information Security*, May 2004.

[8] P. Dasgupta. Trust as a commodity. In D. Gambetta, editor, *Trust: Making and Breaking Cooperative Relations*. Basil Blackwell, 1988.

[9] P. Dasgupta. Social capital and economic performance: Analytics. In E. Ostrom and T. Ahn, editors, *Foundations of Social Capital*. Edward Elgar, 2003.

[10] J. Feigenbaum, C. Papadimitriou, R. Sami, and S. Shenker. A BGP-based mechanism for lowest-cost routing. In *PODC '02: Proceedings of the twenty-first annual symposium on Principles of distributed computing*, pages 173–182. ACM Press, 2002.

[11] M. Feldman and J. Chuang. Hidden-action in multi-hop routing. In *Second Workshop on the Economics of Peer-to-Peer Systems*, June 2004.

[12] M. Feldman, K. Lai, I. Stoica, and J. Chuang. Robust incentive techniques for peer-to-peer networks. In *EC '04: Proceedings of the 5th ACM conference on Electronic commerce*, pages 102–111. ACM Press, 2004.

[13] M. Feldman, C. Papadimitriou, J. Chuang, and I. Stoica. Free-riding and whitewashing in peer-to-peer systems. In *Proceedings of the Third Annual Workshop on the Economics of Information Security*, 2004.

[14] D. Goodhart. Too diverse? *Prospect*, February 2004.

[15] R. Hinde and J. Groebel. *Cooperation and Pro-Social Behaviour*. Cambridge University Press, 1991.

[16] S. D. Kamvar, M. T. Schlosser, and H. Garcia-Molina. The eigentrust algorithm for reputation management in P2P networks. In *WWW '03: Proceedings of the twelfth international conference on World Wide Web*, pages 640–651. ACM Press, 2003.

[17] T.-W. Ngan, D. S. Wallach, and P. Druschel. Enforcing fair sharing of peer-to-peer resources. In *Proceedings of the Second International Workshop on Peer-to-Peer Systems*, March 2003.

[18] N. Nisan. Algorithms for selfish agents: Mechanism design for distributed computation. *Lecture Notes in Computer Science*, 1563:1–15, 1999.

[19] E. Ostrom and T. Ahn, editors. *Foundations of Social Capital*. Edward Elgar, 2003.

[20] W. Powell and P. Brantley. Competitive cooperation in biotechnology: Learning through networks? In N. Nohria and R. Eccles, editors, *Networks and Organizations*. Harvard University Press, 1992.

[21] A. Przeworski. *Democracy and the Market*. Cambridge University Press, 1991.

[22] S. Schecter. Toward econometric models of the security risk from remote attacks. In *Proceedings of the Third Annual Workshop on the Economics of Information Security*, 2004.

[23] I. Serageldin and P. Dasupta, editors. *Social Capital: A Multifaceted Perspective*. World Bank, 2001.

[24] A. Serjantov and R. Anderson. On dealing with adversaries fairly. In *Proceedings of the Third Annual Workshop on Economics and Information Security*, May 2004.

[25] J. Shneidman and D. C. Parkes. Rationality and self-interest in peer to peer networks. In *Proceedings of the Second International Workshop on Peer-to-Peer Systems*, March 2003.

[26] I. Stoica, R. Morris, D. Karger, M. F. Kaashoek, and H. Balakrishnan. Chord: A scalable peer-to-peer lookup service for internet applications. In *Proceedings of the ACM SIGCOMM '01 Conference*, San Diego, California, August 2001.

[27] H. R. Varian. Economic mechanism design for computerized agents. In *Proceedings of USENIX Workshop on Electronic Commerce*, 1995.

[28] M. Yunus. Group-based savings and credit for the rural poor. In *United Nations Inter-Agency Panel on People's Participation*, 1983.