Theory and Decision (2009) 67:101–122
DOI 10.1007/s11238-007-9089-4

HANNU VARTIAINEN

A SIMPLE MODEL OF SECURE PUBLIC COMMUNICATION

ABSTRACT. Public communication is *secure* if a hostile third-party cannot decode the messages exchanged by the communicating parties. In Nash equilibrium, communication by computationally unbounded players cannot be secure. We assume complexity averse players, and show that a simple, secure, and costless communication protocol becomes available as the marginal complexity cost tends to zero.

KEY WORDS: secure communication, complexity costs, trading

JEL CLASSIFICATION. D02, C72

1. INTRODUCTION

Security is instrumental for well functioning markets. Consider the following canonical scenario: A beneficial transaction is available for two traders, a "team", but an intruder may intercept and steal the surplus. To avoid interception, the team members need to agree *secretly* how to execute the transaction. Communication through which such agreement is reached is called *secure*.

What if only *public* communication channels are available for the team? Then security requires that messages are translated— *encrypted*—in a fashion that translating them back— *decrypting*—is prohibitively expensive to all but the desired party. The question is how to design a protocol that is more costly to decrypt than to encrypt.

Modern cryptographics literature relies on the assumption that certain type of computational tasks are more demanding than others. This assumption permits protocols, i.e., functions,

to have the *trapdoor* property: computable in a reasonable time to one direction but to the other only with the right *key* (the "trapdoor"). While most computer scientists believe that trapdoor functions do exist, the assumption is difficult, perhaps impossible, to verify.[1]

The above canonical model of secure communication is a strategic scenario *par excellence*, calling for game theoretic analysis. However, the assumption of the existence of trapdoor functions is in conflict with the key paradigms of game theory; that the framework is common knowledge and that the players are rational. The very notion of Nash equilibrium rules out the possibility that public communication conveys information to some party but not to the other. Hence, the game theory has had surprisingly little to say *about* public key cryptographics (for related literature, see the discussion in the end of this introduction).

The aim of this article is to analyze secure communication in the language of there. Our epistemological assumptions come from there: the framework is common knowledge and there are no limitations on players' ability to understand, compute or solve problems. *No* trapdoor functions are used. The crucial assumption concerns *preferences*. They reflect aversion towards strategic complexity. Hence a player faces a trade-off between his communicative goal and the simplicity of his strategy. Standard game theoretic tools remain available to analyze secure communication.

We study the canonical game discussed in the first paragraph: (1) First both team members release public messages or keys (natural numbers). (2) Conditional on messages, each player chooses an action (a natural number). Team members are rewarded if they choose the same action which is different from the intruder's action, but only the intruder is rewarded if all players choose the same action.

Complexity is modeled in reduced form. As a proxy of the complexity of a strategy, we take the *number of its contingencies*.[2] Complexity aversion manifests itself in the assumption that more complex strategies are more costly to implement. Importantly, we assume that the complexity cost is

equal across the players. Hence, differences in computational capacity do not play any role in the analysis.

We show that a natural, simple, and secure communication protocol emerges in Nash equilibrium.[3] Let us sketch the protocol. There are n^2 actions the team members may choose. Actions are factored into n "rows" and n "columns", and one of the team members, say A, is let to announce one of the rows and the other, B, one of the columns. The action that both team members choose is then the intersection of the announcements. Crucially, since A knows which row he is about to choose, he only needs to make his action contingent on the n columns that B might choose (and vice versa for player B), to choose the right action. Hence, each team member invests in n contingencies. However, to guarantee that he will know the action of the team, the intruder needs to understand what each combination of messages implies which means that he has to invest in n^2 contingencies. Thus the total cost of the team members is $n+n$ contingencies whereas for the intruder, it is $n \cdot n$ contingencies. For $n > 2$, the cost of eavesdropping is higher than coordination.

One expects the impact of complexity to go down as the complexity cost, i.e., the marginal cost of a contingency, decreases. An important question is what happens in the limit, when the complexity neutral preferences are reached. We show that then not only does the intruder's success probability becomes negligible but also the team members' (combined) communication costs tend to zero. Moreover, for low enough complexity cost, the incentive compatibility of the ciphering protocol is guaranteed—no team member wants to free-ride at the expense of the other team member's effort.

We conclude that there is an important discontinuity in the equilibrium correspondence. When the marginal complexity cost tends to zero, perfectly secure and costless communication becomes feasible whereas under zero cost, all communication is totally insecure. Thus, complexity-neutral behavior cannot even be approximated in a model with computational constraints.

1.1. *Related literature*

A successful coordination by the team can be interpreted as a *correlated* equilibrium of the canonical game. A related literature considers the question of whether all correlated equilibria can be generated as the equilibrium outcome of a pure communication process among the players. In other words, whether players can coordinate their actions via pre-play communication without letting them know too much on the others' private information or intentions. Forges (1986) and Bàràny (1992) identify communication protocols achieving this desideratum, and Gossner (1998b) generalizes the notion of a secure protocol. Other recent advances in this literature include Ben-Porath (1998), Gerardi (2004), Lehrer (1996), and Lehrer and Sorin (1997). Izmalkov et al. (2005) show that a simple ballot-box mechanism securely implements any choice rule, including correlated equilibria.

This literature is not concerned with the possibility of achieving coordination through *public* communication, which is the theme of this article.[4] The only exception on public communication (that I am aware of) are Gossner (1998a) and Urbano and Vila (2002), who study how to apply cryptographic methods and boundedly rational agents to obtain correlation. While this is related to our question, the important difference is that we do not assume trap-door functions. Computational bounds are well-specified, and only elementary techniques are used. As we concentrate on a specific setup, sharp results can be achieved concerning the limit behavior as complexity costs tend to zero.

First, we define the setup and discuss the benchmark case of complexity neutral agents. Then we introduce complexity costs and the ciphering protocol, and show that the protocol becomes more powerful as the complexity costs go down. All proofs and one middle-result are relegated to the appendix.

2. THE SETUP

There are three players, team members Alice (A) and Bob (B), and an intruder Eve (E), each choosing an action from

an *infinite* set X.[5] A and B engage in a mutually beneficial transaction if they commit the same action that is different than E's action. If all players choose the same action, then E gets a prize while A and B receive no surplus. In all other cases, no one receives surplus.

More formally, the vNM payoff functions u_A, u_B, and u_E of the players are as follows: for actions $x_i \in X$ of team members $i = A, B$, and $x_E \in X$ of the intruder,

$$u_A[x_A, x_B, x_E] = \begin{cases} R_A > 0, & \text{if } x_A = x_B \neq x_E, \\ 0, & \text{otherwise,} \end{cases}$$

$$u_B[x_A, x_B, x_E] = \begin{cases} R_B > 0, & \text{if } x_A = x_B \neq x_E, \\ 0, & \text{otherwise,} \end{cases}$$

$$u_E[x_A, x_B, x_E] = \begin{cases} R_E > 0, & \text{if } x_A = x_B = x_E, \\ 0, & \text{otherwise.} \end{cases}$$

Action triple $(x_A, x_B, x_E) \in X \times X \times X$ constitutes a pure strategy *Nash equilibrium* if and only if

$$\begin{aligned} u_A[x_A, x_B, x_E] &\geq u_A[x'_A, x_B, x_E], \text{ for all } x'_A \in X. \\ u_B[x_A, x_B, x_E] &\geq u_B[x_A, x'_B, x_E], \text{ for all } x'_B \in X, \\ u_E[x_A, x_B, x_E] &\geq u_E[x_A, x_B, x'_E], \text{ for all } x'_E \in X. \end{aligned} \tag{1}$$

The intruder can do at least as well as $i \in \{A, B\}$ by simply imitating i's strategy x_i. Thus, in any pure strategy, Nash equilibrium the team members get zero payoff (in some equilibria the intruder gets payoff 1).[6]

Note, however, that there is a desirable *correlated* equilibrium where the team members successfully coordinate their random action over X, and the intruder intercepts with a negligible probability. When correlation needs to be conducted via cheap talk (without a mediator), the team's success depends on the publicity of the messages. With *private* messages, i.e., messages that are observable only to the team members, the desired correlated equilibrium can be achieved. More interesting—and relevant—is the case of *public* communication. In this case, the desired correlated equilibrium cannot be achieved.

For suppose that in equilibrium the team members condition their actions on the observed messages. Since the messages are public, and strategies mutually known, the post-message situation is not different from (1). Thus, by imitating a team member i's message contingent action, the intruder is able to do at least as well as i and, hence, team coordination without the interception is not feasible. No matter how sophisticated communication protocol the team members come up with, they cannot fool a rational intruder.

We say that public communication is *secure* if the team members choose the same action that is different from the intruder's action with probability one.

CLAIM 1. *Secure public communication is not feasible in Nash equilibrium.*

3. SECURE PUBLIC COMMUNICATION

3.1. *Complexity considerations*

This section develops a model of public communication under complexity aversion.[7] The order of moves is: 1. Team members send costless public messages in $\mathbb{N}$. 2. All players choose their actions in X.

Functions from $\mathbb{N} \times \mathbb{N}$ to X are called message contingent *action plans*. Denote by $x_A(\cdot), x_B(\cdot)$, and $x_E(\cdot)$ the message contingent action plans of A, B, and E, respectively.

Strategies are defined as follows: Team member $i = A, B$ chooses a probability distribution μ_i over the set of messages $m_i \in \mathbb{N}$. Any player $i = A, B, E$ chooses a probability distribution ξ_i over the set of his action plans $\{x_i : \mathbb{N} \times \mathbb{N} \to X\}$ (with countable support, for simplicity). The chosen action plan x_i implements choice $x_i(m_A, m_B)$ where m_A and m_B are the realized messages of A and B, respectively.

Increasing the complexity of the action plan x_i is now costly. As the proxy of the complexity of a strategy, we take the *number of its contingencies*. Strategy that contains more contingencies is more costly to implement.

Formally, let $x:\mathbb{N}\times\mathbb{N}\to X$ be an action plan. Given an outcome $z\in X$, define the preimage $x^{-1}(z)=\{m\in\mathbb{N}\times\mathbb{N}: x(m)=z\}$. Function x^{-1} spans a partition on $\mathbb{N}\times\mathbb{N}$. For any $Z\subset X$, denote the cardinality of $x^{-1}(Z)$ by $\left|x^{-1}(Z)\right|$. Now $\left|x^{-1}(x(\mathbb{N}\times\mathbb{N}))\right|$, or simply $\left|x^{-1}\right|$, is the measure of the number of contingencies in x.[8] We let $\left|x^{-1}\right|=0$ to mean that a player is inactive (and obtains zero payoff).

The complexity cost is captured by a function $c:\mathbb{N}\to\mathbb{R}_+$, where $c(k)$ is the cost of implementing an action plan with k contingencies. We assume that $c(k)-c(k-1)$ is non-decreasing in $k=1,2,\ldots,$ and that $c(1)>c(0)=0$. In particular, we assume that c is *equal* across players.[9]

Players' expected payoffs are

$$E_{\xi_A,\xi_B,\xi_E,\mu_A,\mu_B}\left\{u_A[x_A(m),x_B(m),x_E(m)]-c(\left|x_A^{-1}\right|)\right\},$$
$$E_{\xi_A,\xi_B,\xi_E,\mu_A,\mu_B}\left\{u_B[x_A(m),x_B(m),x_E(m)]-c(\left|x_B^{-1}\right|)\right\},$$
$$E_{\xi_A,\xi_B,\xi_E,\mu_A,\mu_B}\left\{u_E[x_A(m),x_B(m),x_E(m)]-c(\left|x_E^{-1}\right|)\right\}.$$

Each player wants to maximize his expected gain from real actions but is also concerned about the complexity of his action plan.

While for any given μ_A,μ_B, the trade-off between the goals seems, at first, symmetric across players, it is not since a team member i can *internally coordinate* μ_i and ξ_i by choosing $x_i(\cdot)$ conditional on m_i that he is about to send. To capture this formally, we specify a family of functions $\{x_i^k\}_{k\in\mathbb{N}}$, that are i's, $i=A,B$, message contingent action plans conditional on his message k. We may then rewrite the players' expected payoffs

$$E_{\xi_E,\mu_A,\mu_B}\left\{u_A[x_A^{m_A}(m_B),x_B^{m_B}(m_A),x_E(m)]-c(\left|(x_A^{m_A})^{-1}\right|)\right\},$$
$$E_{\xi_E,\mu_A,\mu_B}\left\{u_B[x_A^{m_A}(m_B),x_B^{m_B}(m_A),x_E(m)]-c(\left|(x_B^{m_B})^{-1}\right|)\right\},$$
$$E_{\xi_E,\mu_A,\mu_B}\left\{u_E[x_A^{m_A}(m_B),x_B^{m_B}(m_A),x_E(m)]-c(\left|x_E^{-1}\right|)\right\}.$$

Consider the following simple *ciphering protocol* (or protocol for short). Let y be an injection from $\mathbb{N}\times\mathbb{N}$ to X (recall that X is infinite).

DEFINITION 2. *The team engages in the* n-protocol *if* μ_A *and* μ_B *are uniform distributions on* $\{1, \ldots, n\}$, *and* $x_A^{m_A}(m_B) = x_B^{m_B}(m_A) = y(m_A, m_B)$, *for all* $(m_A, m_B) \in \{1, \ldots, n\}^2$.[10]

Under the n-protocol, the team members send random, independent messages m_A and m_B from the n-element set. After observing the messages, they choose the actions according to plan y. Since y is an injection, the number of contingencies in y is n^2.[11] Note, however, that $x_i^{m_i}$, which is an injection from $\{1, \ldots, n\}$ to X, contains only n contingencies. Thus i's cost of following the n-protocol is only $c(n)$.

We say that the intruder *breaks* the n-protocol at m if $x_E(m) = y(m)$. To break the n-protocol under all $m \in \{1, \ldots, n\}^2$, the intruder has to invest in n^2 contingencies. But this strategy may not be optimal for the intruder. The next lemma characterizes intruder's optimal investment in contingencies or, equivalently, optimal probability of intervening.

LEMMA 3. *Let the team engage in the* n*-protocol, and let* E*'s best response be* x_E. *Then there is* k *and* $M \subseteq \{1, \ldots, n\}^2$ *with* $|M| = k - 1$ *such that* $x_E(m) = y(m)$ *for all* $m \in M$ *and* $x_E(m) \in y(M)$ *for all* $m \notin M$. *Moreover,*

$$k = \begin{cases} n^2 & \text{implies} \quad 1/n^2 \geq c(n^2) - c(n^2 - 1), \\ 1, \ldots, n^2 - 1 & \text{implies} \quad c(k+1) - c(k) \\ & \qquad\qquad \geq 1/n^2 \geq c(k) - c(k-1), \\ 0 & \text{implies} \quad c(1) \geq 1/n^2. \end{cases}$$

That is, at the optimum, the intruder invests in k contingencies such that whenever a message from a set M with cardinality $k - 1$ materializes, he knows which action the team is about to take. After seeing any of the remaining $n^2 - k + 1$ pairs of messages, he only knows that the message pair was not from the set M. Finally, the optimal number of contingencies for the intruder depends only on the marginal cost of a contingency and the probability of investing to a priori correct message, i.e., breaking the protocol at m (which is constant over messages since $\mu_A \mu_B$ is uniform over $\{1, \ldots, n\}^2$).

Denoting the k in Lemma 3 as $k(n)$, note that the intruder breaks the n-protocol with a priori probability $k(n)/n^2$. Since the differences in $c(\cdot)$ are non-decreasing, it is immediately clear that as n becomes large, $k(n)$ falls and, in the limit, approaches 0. Hence increasing the size of the ciphering protocol decreases the probability that the protocol will be broken.

3.2. *Increase in computational capacity*

We now study how evolution of strategic complexity costs—which we interpret as evolution of players' computational capacity—affects their behavior. We allow players computational capacity to increase. They, therefore, become "more rational".

Assume that there is a continuously differentiable, convex function $\psi : \mathbb{R}_+ \to \mathbb{R}_+$ that agrees with c on $\mathbb{N}$, i.e., $c(k) = \psi(k)$ for all $k \in \mathbb{N}$. Function $\psi(k)$ reflects the cost of computing a block ("byte") of computational units ("bits"). Let the size of a computable block increase in time $t = 0, 1, \ldots$ such that $c = c^0$ and period t cost function c^t satisfies

$$c^t(k) = \psi\left(\frac{k}{4^t}\right), \text{ for all } k \in \mathbb{N}.$$

Then $c(k) = c^t(4^t k)$ for all t, for all k. One could think each period t as a technology generation that quadruples players' computational capacity, i.e., the amount of contingencies a player can compute with a given cost.[12] (Quadrupling is for expositional convenience.)

Since ψ is a non-decreasing function, $c(k) \geq c^t(k)$ for all $t \in \mathbb{N}$ and $\lim_{t\to\infty} c^t(k) = \lim_{t\to\infty} \psi(k/4^t) = 0$, for all $k \in \mathbb{N}$.

Now we find a rule that tells the team which n-protocol to choose under each t. Let the team members increase the size of the protocol as t increases according to scheme $n : \mathbb{N} \to \mathbb{N}$ such that

$$n(t) = t2^t, \text{ for all } t \in \mathbb{N}. \tag{2}$$

Thus $n(\cdot)$ suggests a path of evolution of a ciphering protocol when the computation capacity increases.

LEMMA 4. *Let* $n(\cdot)$ *satisfy* (2). *For any* $\varepsilon > 0$, *there is* $t^{\varepsilon} \in \mathbb{N}$ *such that for all* $t \geq t^{\varepsilon}$, *the* $n(t)$*-protocol is broken by the intruder with probability less than* ε.

Rule (2) drives up the amount of ciphered information faster than the intruder's information processing capacity improves. After some point, a payoff maximizing intruder reduces the ratio of information he processes. As a consequence, the intruder is able to break the $n(t)$-protocol less often. In the limit, he never breaks the protocol.

However, the $n(t)$-protocol requires also the team members to process more information as t increases. We now argue that beyond some level of technological development, the increased computational costs are overshadowed by the increased benefits from less likely interception by the intruder. Hence, technology improvement plays in favor of the team members.

PROPOSITION 5. *Let* $n(\cdot)$ *satisfy* (2). *For any* $\varepsilon > 0$, *there is* $t^{\varepsilon} \in \mathbb{N}$ *such that, for all* $t \geq t^{\varepsilon}$, *the* $n(t)$*-protocol generates a team member* $i = A, B$ *a surplus of at least* $R_i - \varepsilon$.

Hence, as the information processing capacity of the players increases, the team members' information processing advantage over that of the intruder's increases. In the limit, the team can share costlessly a secret that is prohibitively costly for the intruder to break.

Note that, since there are no *fixed costs* in participating the game, intervening with positive probability is always profitable for the intruder. In that sense, the $n(t)$-protocol induces secure communication only in the probabilistic sense: as t becomes large, the interception probability goes to zero. However, if there is even a small fixed cost associated to participating the game, e.g., fixed investment into computational capacity, then there is high enough t such that the intercept will never intervene, since the payoff from doing so approaches zero as t becomes large. Hence, *full security* will be reached after some t under $n(t)$-ciphering.

3.3. *Incentive compatibility*

A potential weakness of Proposition 5 is that it assumes the team members to *commit* to the ciphering strategy. This would not be a problem if a cipher could be contracted upon, or if the team members behavior can be disciplined via reputation. However, since anonymity is a characteristic feature in the markets where security is needed, such devices might not be available. A priori, hence, it might not be privately optimal for a team member to invest into all contingencies required by the protocol. The cost of not doing so spreads over the whole team whereas the savings are private.[13]

Since free-riding would render the protocol unworkable, it is important to make sure that it is *also* privately optimal, or *incentive compatible*, for the team members to obey the rules of the n-protocol, given the intruder's best response strategy.

We now argue that for high enough t, incentive compatibility of the $n(t)$-protocol is guaranteed. Assume that the intruder adheres to a best response ξ_E^t that randomizes uniformly over *all* best response action plans x_E against the $n(t)$-protocol.[14] Hence, as in Lemma 3, letting $k(n(t))$ be the optimal number of messages broken by the intruder, it follows that the a priori probability $p(t)$ of breaking any fixed m is $p(t)=k(n(t))/n(t)^2$. Since $n(t)$ tends to infinity in t and, as discussed in the end of Section 3.1, $k(n)$ tends to zero in n, the probability $p(t)$ tends to zero in t. Finally, it follows (see Lemma 7 in the appendix) that team member i finds it privately optimal to obey the $n(t)$-protocol if

$$c^t(n(t)) - c^t(n(t)-1) \leq R_i\left(\frac{1-p(t)}{n(t)}\right). \tag{3}$$

For incentive compatibility of the $n(t)$-protocol under t, it therefore suffices to show that (3) holds. The next proposition shows that it does for high enough t. Thus for high enough t, neither team member wants to free ride.

PROPOSITION 6. *Let $n(\cdot)$ satisfy (2). There is t^* such that, for all $t \geq t^*$, the $n(t)$-protocol is incentive compatible.*

In other words, for high enough t, the $n(t)$-protocol is a team member's best response given that the other team member obeys his $n(t)$-protocol strategy, and the intruder plays ξ_E^t.

From Propositions 5 and 6, it follows that for any $\varepsilon > 0$ there is $t^\varepsilon \in \mathbb{N}$ such that, for all $t \geq t^\varepsilon$, the $n(t)$-protocol is incentive-compatible and generates a team member i a surplus of at least $R_i - \varepsilon$.

3.4. *Example: The linear case*

For a concrete example, consider the case $R_A = R_B = R_E = 1$ with *linear* complexity cost $c(n) = \bar{c}n$ for some $\bar{c} \geq 0$. Recall that under "full" computational capacity, i.e., when $\bar{c} = 0$, secure communication is never feasible (Claim 1) and the surplus to the team members is at most 1/4 (see Endnote 3). Paradoxically, however, when $\bar{c}$ is small and strictly positive, secure communication is guaranteed: since $c(k+1) - c(k) = \bar{c}$ for all k, it follows by Lemmas 3 and 7 (see appendix), that the n-protocol is incentive compatible and never broken by the intruder if $n \geq \bar{c} \geq 1/n^2$.

We now argue that the equilibrium payoff correspondence is discontinuous. Note that for the team, it suffices to invest in the minimal feasible protocol size $n(\bar{c}) = \min\{n : n^2 \geq 1/\bar{c}\}$. Number $n(\bar{c})$ approaches infinity as $\bar{c}$ tends to 0. By construction, the cost of $n(\bar{c})$-protocol to a team member is $n(\bar{c})\bar{c}$. Since communication is secure for small strictly positive $\bar{c}$, the team members gain surplus $1 - n(\bar{c})\bar{c}$. By the construction of $n(\bar{c})$, we have $n(\bar{c})\bar{c} \leq n(\bar{c})/(n(\bar{c}) - 1)^2$. Since the right-hand side approaches zero as $\bar{c} \to 0$, it follows that *the equilibrium payoffs of a team member tend to 1 and hence exhibit discontinuity at* $\bar{c} = 0$.

That the equilibrium payoff correspondence can be sensitive to players' information processing costs is observed in many models of strategic complexity considerations. For example, Rubinstein (1986) and Abreu and Rubinstein (1988) show that a *lexicographic* computational cost may have large consequences on equilibria in infinitely repeated games.[15]

4. DISCUSSION

The n-protocol developed in this article gives n over n^2 advantage to legitimate communicators relative to any intruder. While this advantage is not huge, and might be somewhat sensitive to the exact form of the players' cost functions under a given n, we have argued that in the limit, when computational costs become small, the difference becomes decisive: the intruder becomes powerless when facing an appropriately designed ciphering protocol. Thus one expects the power usefulness of the protocol to increase alongside with technological development.

The now standard public key cryptographic methods give a much bigger advantage to the legitimate communicators. However, the methods are critically based on the *assumption* that trapdoor functions exist.[16] While most computer scientists believe this assumption is warranted, it is difficult, perhaps impossible, to prove. In particular, the RSA algorithm,[17] from which the many used cryptographic algorithms origin, falls into this category.[18]

However, there is an aspect that increases the attractiveness of the n-protocol relative to trapdoor-based ciphering methods. For example, RSA is based on the assumption that factoring primes is difficult. Large enough factorization problems guarantee security of RSA against any known algorithm. But factoring primes becomes easier as computer hardware and algorithms improve. While the former is not seen as a threat to RSA since one can always increase the key size, advances in the algorithm design could, at least in principle, create a problem of factoring becoming a relatively easier computational operation. However, with n-protocol, which does not employ trapdoor functions, algorithmic (or hardware) advances make no difference. What is needed is that computational capacity is (sufficiently) *symmetrically bounded* across players.

The crucial assumption we make is that randomizing is free (cheap). What makes a team member's position advantageous relative to that of the intruder's is his ability to

coordinate for free, the different dimensions of his actions. Owing to the randomness of this choice, any outsider needs to invest in capacity to condition his action also to the team members' actions. Thus, as long as randomization is costless—the basic paradigm of the game theory—the team member has an advantage over the outsider. In particular, there is no need to assume that some computational tasks are more costly than others. Indeed, we assume that all parties are equally capable of solving *all kinds* of computational problems.

What if randomization is *not* costless?[19] Little is lost if randomization and computation decisions are separate tasks, i.e., if the act of randomization is made before the investment on computational capacity.[20] In that case, it is natural to assume that randomization and computation costs are separately additive entities in the payoff functions. Then, again, the size of combined computational effort of team members grows in the arithmetic order. Comparing this to the geometrically increasing cost of the intruder would generate qualitatively unchanged results.

One could also interpret randomization differently. Following the *purification* argument by Harsanyi (1973), mixed strategies could be thought as the limit of a sequence of pure strategy equilibria of a game with vanishingly small uncertainty about the opponents payoffs. Under this interpretation, the team members could be "born" with their keys (for example a real number whose first $n(t)$ digits they communicate).

An alternative possibility is that the team members economize on computation costs across transactions. Rather than randomizing over the keys before every transaction, the team members could use the same key repeatedly. Then they would invest on randomization only before the first transaction. Of course new reputation and learning questions would emerge but in anonymous markets, where intruders face different team members in every period, they might notedly be of crucial importance.

ACKNOWLEDGEMENTS

I thank Klaus Kultti, Hannu Salonen, and Juuso Välimäki for useful comments.

APPENDIX

Proof of Lemma 3. E's expected payoff is

$$R_E \sum_{m\in\{1,\dots,n\}^2} \Pr\{m\}\Pr\{y(m)=x_E(m)\} - c(\left|x_E^{-1}\right|), \tag{4}$$

where $\Pr\{m\}$ is the probability that $m \in \{1,\dots,n\}^2$ realizes under n-protocol, and $\Pr\{y(m)=x_E(m)\}$ is E's interception probability conditional on message m. The set of messages m' that lead E to choose $y(m)\in\{1,\dots,n^2\}$ is $x_E^{-1}(y(m))$. Since y is a bijection, and m is uniformly distributed on $\{1,\dots,n\}^2$, we have

$$\Pr\{y(m)=x_E(m)\}=\frac{1}{\left|x_E^{-1}(y(m))\right|}.$$

We solve the optimal strategy x_E in two steps. The first step is to identify how E optimally partitions $\{1,\dots,n\}^2$ into k cells. Since $\Pr\{m\}=1/n^2$ under n-protocol, the first element in (4) equals

$$\frac{1}{n^2}\sum_{m\in\{1,\dots,n\}^2}\frac{1}{\left|x_E^{-1}(y(m))\right|}. \tag{5}$$

For any m, m' such that $\left|x_E^{-1}(y(m))\right| \le \left|x_E^{-1}(y(m'))\right|$ it holds that

$$\frac{1}{\left|x_E^{-1}(y(m))\right|}+\frac{1}{\left|x_E^{-1}(y(m'))\right|}<\frac{1}{\left|x_E^{-1}(y(m))\right|-1} + \frac{1}{\left|x_E^{-1}(y(m'))\right|+1}.$$

Therefore, since x_E maximizes (5), x_E^{-1} is a partition of $\{1,\dots,n\}^2$ in which one cannot move one element of a cell

into a bigger cell. This means that x_E^{-1} must contain $k-1$ singleton cells, and one cell with $n^2-(k-1)\geq 1$ elements. Let $M\in x_E^{-1}$ be the set of messages with $|M|=n^2-(k-1)$. Then $\{1,\ldots,n\}^2\setminus M$ contains $k-1$ elements. Given M, (5) can be written

$$\begin{aligned}\frac{1}{n^2}\sum_{m\in\{1,\ldots,n\}^2}\frac{1}{\left|x_E^{-1}(y(m))\right|}&=\frac{1}{n^2}\left(\sum_{m\in M}\frac{1}{|M|}+\sum_{m\notin M}\frac{1}{|\{m\}|}\right)\\&=\frac{1}{n^2}\left(1+(k-1)\right)\\&=\frac{k}{n^2}.\end{aligned}$$

Second, since $k-1$ is the optimal number of singleton cells, we have that, for any $k'=1,\ldots,n^2$,

$$\frac{R_E k}{n^2}-c(k)\geq\frac{R_E k'}{n^2}-c(k').$$

Equivalently

$$R_E\left(\frac{k-k'}{n^2}\right)\geq c(k)-c(k')$$

Since $c(k+1)-c(k)$ is nondecreasing in k, it suffices to check for all $k'=k+1\leq n^2$ and and for all $k'=k-1\geq 0$. These considerations give the result. □

Proof of Lemma 4. For any $t\in\mathbb{N}$, find, if possible, the maximal $k^t\in\mathbb{N}$ such that

$$\frac{R_E}{n(t)^2}\geq c^t(k^t)-c^t(k^t-1). \tag{6}$$

That is,

$$\frac{R_E}{n(t)^2}\geq\psi\left(\frac{k^t}{2^{2t}}\right)-\psi\left(\frac{k^t-1}{2^{2t}}\right),$$

Since ψ is convex and differentiable, we have

$$\frac{R_E}{n(t)^2}\geq\psi'\left(\frac{k^t}{2^{2t}}-\frac{1}{2^{2t}}\right)\frac{1}{2^{2t}}. \tag{7}$$

Since $n(\cdot)$ satisfies (2) we have, by (7),

$$\frac{R_E}{t^2} \geq \psi' \left(\frac{k^t}{2^{2t}} - \frac{1}{2^{2t}} \right). \tag{8}$$

Let x_E be E's best response to the $n(t)$-protocol. The probability that E breaks the $n(t)$-protocol is $p(t) = \left| x_E^{-1} \right| / n(t)^2$. If $\left| x_E^{-1} \right| = 0$, then the claim follows. Assume that $\left| x_E^{-1} \right| > 0$. Then, by Lemma 3, $k^t = \left| x_E^{-1} \right|$ where k^t is defined as in (6). We have, by (2),

$$p(t) = \frac{k^t}{t^2 2^{2t}}.$$

By (8),

$$\frac{R_E}{t^2} \geq \psi' \left(t^2 p(t) - \frac{1}{2^{2t}} \right).$$

Since R_E is a number, and since ψ' is a nondecreasing and continuous function, it follows that the maximum $p(t)$ that meets the inequality must tend to zero as t becomes high. □

Proof of Proposition 5. Let n satisfy (2). Then

$$c^t\,(n(t)) = c^t(\lambda 2^t) = \psi \left(\frac{t}{2^t} \right).$$

Now $t/2^t$ tends to 0 as t becomes large. Hence, by continuity of ψ, also $c^t\,(n(t))$ tends to zero in t. Under $n(t)$-protocol, a team member i's payoff is

$$(1 - p(t))\, R_i - c^t(n(t)), \tag{9}$$

where $p(t)$ is the probability that E breaks the $n(t)$-protocol. Since $c^t\,(n(t))$ tends to 0 and by Lemma 4, the first term in (9) tends to R_i as t becomes large, the proposition is established. □

LEMMA 7. *Let team member j engage in the n-protocol. Suppose that all messages m are broken with probability p. Then team member i's best response x_i satisfies $x_i(m) \neq y(m)$ for some m only if $c(n) - c(n-1) > (1-p) R_i / n$.*

Proof. The proof proceeds along the lines that of Lemma 3. Since all $m \in \{1, \ldots, n\}^2$ are intervened with probability p, i's best response x_i maximizes i's expected payoff

$$R_i \Pr\{x_i(m_A, m_B) = y(m_A, m_B) \neq x_E(m_A, m_B)\} - c^t(|x_i^{-1}(y(\{m_i\} \times \mathbb{N}))|). \tag{10}$$

Now, for any m_j,

$$\Pr\{y(m_A, m_B) = x_i(m_A, m_B) \neq x_E(m_A, m_B)\} = \frac{1-p}{\left|x_i^{-1}(y(m_A, m_B))\right|}.$$

Therefore, the first term in (10) is equal to

$$R \sum_{m_{-i} \in \{1,\ldots,n\}} \frac{\Pr\{y(m_A, m_B) = x_i(m_A, m_B) \neq x_E(m_A, m_B)\}}{n} = R_i \left(\frac{1-p}{n}\right) \sum_{m_{-i} \in \{1,\ldots,n\}} \frac{1}{\left|x_i^{-1}(y(m_A, m_B))\right|}. \tag{11}$$

Assume that $\left|x_i^{-1}(y(\{m_i\} \times \mathbb{N}))\right| = k$. Recall that for any $x, x' \in X$ such that $\left|x_i^{-1}(x)\right| \leq \left|x_i^{-1}(x')\right|$,

$$\frac{1}{\left|x_i^{-1}(x)\right|} + \frac{1}{\left|x_i^{-1}(x')\right|} < \frac{1}{\left|x_i^{-1}(x)\right| - 1} + \frac{1}{\left|x_i^{-1}(x')\right| + 1}.$$

Hence, since x_i maximizes (10), x_i^{-1} spans a partition on $\{1, \ldots, n\}$ where at least all but one cell are singleton sets. Let the number of singleton cells be $k-1$. Let $M \in x_E^{-1}$ be the set of messages with $|M| = n - (k-1)$. Then $\{1, \ldots, n\} \setminus M$ contains $k-1$ elements. Given M, we have

$$\sum_{m_{-i} \in \{1,\ldots,n\}} \frac{1}{\left|x_i^{-1}(y(m_A, m_B))\right|} = \left(\sum_{m \in M} \frac{1}{|M|} + \sum_{m \notin M} \frac{1}{|\{m\}|}\right) = k.$$

Thus, by (11), the first term in (10) becomes

$$R_i \Pr\{x_i(m_A, m_B) = y(m_A, m_B) \neq x_E(m_A, m_B)\}\} = \frac{(1-p) R_i k}{n}.$$

Since x_i is i's best response, and contains x contingencies, we have

$$\frac{(1-p)R_i k}{n} - c(k) \geq \frac{(1-p)R_i k'}{n} - c(k'), \text{ for all } k' \in \mathbb{N}. \quad (12)$$

We ask when this holds for $k=n$. Since the differences in $c(\cdot)$ are increasing, it suffices to check (12) for $k=n-1$. Thus we need

$$\frac{(1-p)R_i}{n} \geq c(n) - c(n-1). \qquad \square$$

Proof of Proposition 6. We construct the following equilibrium: the team members choose $n(t)$-protocol as in (2), and E chooses uniform ξ_E on all x_E's such that $\left|x_E^{-1}\right| / n(t)^2 = p(t)$, where x_E satisfies Lemma 3. Then, by Lemma 7, $n(t)$-protocol is a team member i's best response if and only if

$$c^t(n(t)) - c^t(n(t)-1) \leq R_i \frac{1-p(t)}{n(t)}.$$

By (2), this translates to

$$\psi\left(\frac{t}{2^t}\right) - \psi\left(\frac{t}{2^t} - \frac{1}{2^{2\lambda}}\right) \leq R_i \frac{1-p(t)}{t2^t},$$

or

$$\frac{\psi\left(\frac{t}{2^t}\right) - \psi\left(\frac{t}{2^t} - \frac{1}{2^{2t}}\right)}{\frac{1}{2^{2\lambda}}} \leq R_i \frac{2^t}{t}(1-p(t)). \quad (13)$$

We argue that for high enough t (13) is met. Combining this with Proposition 5 establishes the proof.

Proof that (13) is met for high enough t: By continuity of ψ, the left-hand side converges to

$$\begin{aligned}
&\lim_{t\to\infty} \frac{\psi\left(\frac{t}{2^t}\right) - \psi\left(\frac{t}{2^t} - \frac{1}{2^{2t}}\right)}{\frac{1}{2^{2\lambda}}} \\
&= \lim_{t\to\infty} \left(\lim_{\Delta\to 0} \frac{\psi\left(\frac{t}{2^t}\right) - \psi\left(\frac{t}{2^t} - \Delta\right)}{\Delta} \right) \\
&= \lim_{t\to\infty} \psi'\left(\frac{t}{2^t}\right) \\
&= \psi'(0).
\end{aligned}$$

Since ψ is differentiable and continuous, necessarily $\psi'(0) < \infty$. By Proposition 4, $p(t)$ tends to zero when t become large. Thus, after some threshold, the right-hand side of (13) increases in t and, in the limit, approaches infinity. Hence there is high enough t under which (13) is met. □

NOTES

1. See the discussion in the last section.
2. Dye (1985) makes a similar assumption in a seminal paper. Contingency costs may be due to, say, implemention of a strategy via costly automatae. This approach was pioneered by Rubinstein (1986) and Abreu and Rubinstein (1988). Kalai and Stanford (1988) and Osborne and Rubinstein, (1994) are introductions.
3. This protocol turns out to be formally close to one of the first public key cryptosystems in the literature, the *Merkle's Puzzles,* suggested by Merkle (1974, 1978). See also Diffie (1988).
4. Lehrer and Sorin (1997) consider a protocol where players sends private messages to a mediator who returns with a public (deterministic) announcement.
5. X could be interpreted as an index set of different *languages* that A and B use to exhange valuable information.
6. In the traders' optimal mixed strategy equilibrium all players randomize uniformly on a pair of outcomes. A trader i's payoff is $R_i(1/2)(1-1/2) = R_i/4$, and the intruder's payoff is $R_E(1/2)(1/2) = R_E/4$.
7. The underlying motivation could be that strategies are executed via costly automatae, see e.g. Rubinstein (1998).
8. Equivalently, define a *finite automaton* (τ, Q, f), where Q is the set of states, $\tau : \mathbb{N} \times \mathbb{N} \to Q$ is the transition function, specifying the final state for each message, and $f : Q \to A$ is the outcome function. Automaton now *implements* plan $g(\cdot)$ if $f(\tau(m)) = g(m)$ for all $m \in \mathbb{N} \times \mathbb{N}$. The number of contingencies in g clearly coincides with the minimal number of states of an automaton that implements it. (for introduction, see Rubinstein, 1998, or Osborne and Rubinstein, 1994).
9. Any differences in information processing are not, therefore, due to differences in players' computational abilities.
10. Any element outside $\{1, \ldots, n\}^2$ materializes with zero probability. We assume implictly that if any such message is observed, then the n-protocol associates it to some cell in the partition of $\{1, \ldots, n\}^2$.

11. To be precise at least n^2, depending on whether the zero probability messages are responded by an action in $y(\{1,\ldots,n\}^2)$ or not. We assume that they are, see the previous footnote. Hence n^2 is the appropriate number of contingencies.
12. Moore's law: computational capacity of computers doubles every 12 months.
13. This is a hold-up problem.
14. Note that since there are finitely many bijections from $\{1,\ldots,n(t)\}^2$ to $y(\{1,\ldots,n(t)\}^2)$, the number of the intruder's optimal action plans is finite. Hence ξ_E^t is well-defined.
15. However, in their model computational costs do not *create* new equilibria, which is the key aspect of this article.
16. A trapdoor function is easy to compute to one direction but difficult to the other by anyone not possessing the right key (the "trapdoor"). Thus, the standard algorithms not only assume that computation is costly but also that certain *forms* of computation are inherently more difficult than others.
17. By Rivest et al. 1978.
18. Diffie and Hellman (1976) is the first public key cryptographic algorithm. Also this system is based on the difficulty of factoring primes. Interestingly, Diffie–Hellman uses two-sided keys, as our model does. Surveys of public-key cryptography are given by Diffie (1988), Kaliski (1993) or more extensively, Menezes et al. (1996).
19. A deeply troubling assumption from the viewpoint of game theory.
20. Since the traders have discretion over their own strategies.

REFERENCES

Abreu, D. and A. Rubinstein (1988), The structure of Nash equilibria in repeated games with finite automata, *Econometrica* 56, 1259–1282.

Bàràny, I. (1992), Fair distribution protocols, or how the players replace fortune, *Mathematics of Operations Research* 17, 327–340.

Ben-Porath, E. (1998), Correlation without mediation: expanding the set of equilibrium out-comes by "cheap" pre-play procedures, *Journal of Economic Theory* 80, 108–122.

Diffie, W. and Hellman, M.E. (1976), New directions in cryptrography, *IIE transaction on Information Theory*, 644–654.

Diffie, W. (1988), The first ten years of public-key cryptography, *Proceedings of the IEEE* 76, 560–577.

Dye, R. (1985), Costly contract contingencies, *International Economic Review* 26, 233–250.

Dodis, Y., Halevi, S. and Rabin, T. (2000), A cryptographic solution to a game theoretic problem. *Proceedings of Crypto 2000,* 112–130

Forges, F. (1986), An approach to communciation equilibria, *Econometrica* 58, 1375–1385.
Gerardi, D. (2004) Unmediated communication in games with complete and incomplete information. *Journal of Economic Theory* 114, 104–131.
Gossner, O. (1998a), Repeated games with cryptographically sophisticated players, CORE DP. 9835.
Gossner, O. (1998b), Secure protocols or how communication generates correlation, *Journal of Economic Theory* 83, 69–89.
Harsanyi, J. (1973), Games with randomly disturbed payoffs: a new rationale for mixed-strategy equilibrium points, *International Journal of Game Theory* 2, 1–23.
Izmalkov, S., Lepinski, M. and Micali, S. (2005), Rational secure function evaluationand ideal mechanism design. *Proceedings of FOCS*.
Kaliski B., Jr. (1993), A survey of encryption standards, *IEEE Micro* 13, 74–81.
Kalai, E. and Stanford, W. (1988), Finite rationality and interpersonal complexity in repeated games, *Econometrica* 56, 397–410.
Lehrer, E. (1996), Mediated talk, *International Journal of Game Theory* 25, 177–188.
Lehrer, E. and Sorin, S. (1997), One-shot public mediated talk, *Games and Economic Behavior* 20, 131–148.
Merkle, R. (1978), Secure communications over insecure channels, *Communications of the ACM* 21, 294–299.
Merkle, R. (1974), Secure communications over insecure channels, Ph.D. Thesis, Department of Computer Science, Stanford University.
Menezes, A., van Oorschot, P. and Vanstone, S. (1996), *Handbook of Applied Cryptography*, CRC Press.
Osborne, M. and Rubinstein, A. (1994), *A Course in Game Theory*, MIT Press, Cambridge, MA.
Rivest, R., Shamir, A. and Adleman L. (1978), A method for obtaining digital signatures and public key cryptosystems, *Communications*, 120–126.
Rubinstein, A. (1986), Finite automata play the repeated prisoner's dilemma, *Econometrica* 39, 83–96.
Rubinstein, A. (1998), *Modeling Bounded Rationality*, MIT University Press, Cambridge, MA.
Urbano, A. and Vila, J. (2002), Computational complexity and communication, coordination in two-player games, *Econometrica* 70, 1893–1927.

Address for correspondence: Hannu Vartiainen, Yrjö Jahnsson Foundation, Ludviginkatu 3–5, FIN-00130 Helsinki, Finland E-mail: hannu.vartiainen@yjs.fi
Hannu Vartiainen, Turku School of Economics, Turku, Finland