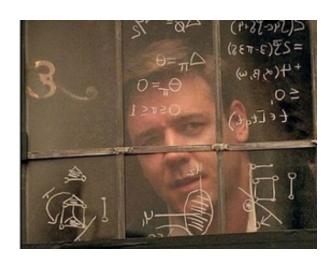
La Teoria dei Giochi per la Sicurezza

Christian Esposito

Dipartimento di Informatica Università di Salerno

esposito@unisa.it





Roadmap

- Introduzione dell'applicazione dei giochi alla sicurezza
- > Esempi di Security Games

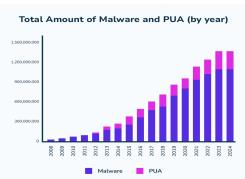






Network Security – 1/12

- Le reti presentano numerosi problemi, tra cui: attacchi, crimini informatici, accesso illegale ai dati, furto di dati, malware, Potentially Unwanted Application (PUA) ecc.
- Gli attacchi alla rete possono causare la perdita di denaro, dati importanti o la reputazione di istituzioni pubbliche o entità private.



I resoconti di nuovi crimini informatici e incidenti indicano che la sicurezza della rete è un argomento impegnativo.





Network Security – 2/12

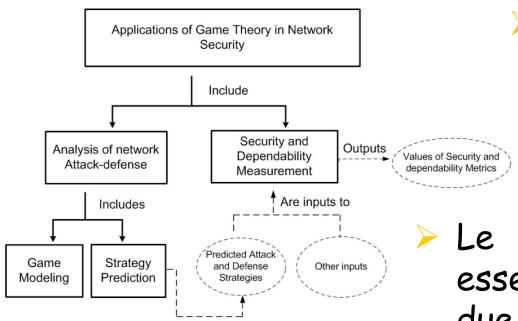
- Le soluzioni tradizionali sono implementate impiegando un dispositivo preventivo, come un firewall, o un dispositivo reattivo, come un programma antivirus, o utilizzandoli insieme.
- > Tali soluzioni hanno delle carenze, e la teoria dei giochi è stata proposta per migliorarle.
 - La debolezza delle soluzioni tradizionali è la mancanza di un quadro decisionale quantitativo.
 - La teoria dei giochi è in grado di analizzare molti possibili scenari con obiettivi opposti prima di determinare il corso appropriato di azioni. Ciò può rendere molto più sofisticato il processo decisionale.





Network Security – 3/12

Nella sicurezza di rete, le interazioni tra aggressori e difensori devono essere modellate in modo da prevedere le azioni di entrambi.



Outputs

Le interazioni tra attaccanti e difensori possono essere modellate come un gioco.

Le applicazioni possono essere classificate in due categorie.





Network Security – 3/12

Nella sicurezza di rete, le interazioni tra aggressori e difensori devono essere modellate in modo da prevedere le azioni di entrambi.

Applications of Game Theory in Network Security

Include

Analysis of network Attack-defense

Includes

Game Modeling

Strategy Prediction

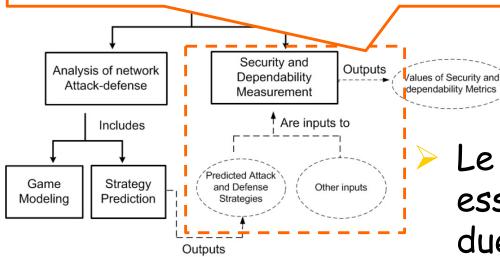
Outputs

<u>> Le interazioni tra</u>

Applicazioni per l'analisi dell'attacco-difesa in rete (decisioni quantitative): l'interazione tra aggressori e difensori come giochi, previsione delle azioni degli aggressori e determinazione della strategia di difesa di risposta.

Network Security – 3/12

Applicazioni per la misurazione della sicurezza e dell'affidabilità della rete: previsione delle strategie degli aggressori e dei difensori e valutazione della sicurezza del sistema in base a questa previsione.



difensori possono essere modellate come un gioco.

Le applicazioni possono essere classificate in due categorie.





Network Security – 4/12

- Le soluzioni tradizionali per la sicurezza di rete mostrano debolezze quando si trovano ad affrontare aggressori sofisticati o ben organizzati.
- > Nello scenario critico dei giochi di attacco-difesa:
 - > Le reti spesso non sono specifiche ma astratte.
 - Lo scenario è un attaccante contro un difensore.
 - Le azioni dell'attaccante sono attaccare o non fare nulla. Le azioni del difensore sono difendersi o non fare nulla.
 - In genere, per modellare il gioco attaccante-difensore viene utilizzato un gioco di Stackelberg.





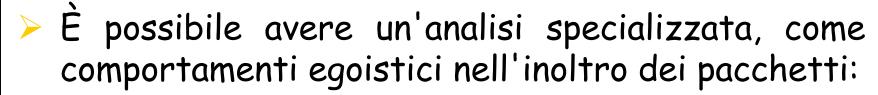
Network Security – 5/12

- Nella maggior parte dei casi, il difensore potrebbe non avere informazioni perfette sul fatto che un nodo della rete sia un aggressore o un utente normale; può solo fare inferenze basate sulla sua convinzione.
- Un modello generico di signal game può essere utilizzato per modellare l'interazione tra ciascuna coppia di nodi costituita da un nodo difensore e uno dei suoi vicini, e viene quindi utilizzato per determinare la migliore strategia di difesa.





Network Security – 6/12



Game model	Features	Solution
Prisoners' Dilemma Game	Introduction of social morality to improve user privacy. Morality state is modeled as a Markov chain.	Nash equilibrium is adopted in an incomplete and a complete information game
Cooperative game	Forming coalitions to enhance a carry-and-forward-based packet-forwarding mechanism	A Nash bargaining formulation to obtain a Pareto-optimal solution
Stochastic game	Transmitting a packet is a Bernoulli process. The game is a repeated asymmetric game with random states.	Punish Only n Times: A distributed algorithm to achieve a Subgame Perfect Equilibrium
Repeated game	A formal-belief-system based on Bayes' rule to revise the other nodes' information under imperfect observation	An iterative update belief algorithm to find the sequential equilibrium
Evolutionary game	Decision-making of nodes based on the limited information of the others. Public Goods game is used to incentivize cooperation.	The strategy of the nodes is updated by comparing their payoff with a randomly chosen neighbor





Network Security – 7/12

Le preoccupazioni relative alla privacy possono essere modellate come giochi.

Privacy issue	Key objective	Game model	Solution
Cryptography	Guarantee that participators keep using the designated service	Two-player static game Dynamic game Extensive form game Repeated game Stochastic two-player zero-sum game	Nash equilibrium Perfect Bayesian equilibrium Bayesian Nash equilibrium Computational Nash equilibrium
Anonymity	Aim to participate in the system without being tracked	Perfect and complete sequential game Repeated game Bayesian game	Subgame Perfect Equilibrium Nash equilibrium Bayesian Nash equilibrium
Information sharing	Protect the privacy information in the information sharing process	Cooperative game Evolutionary game Strategic game Two-player zero-sum Markov game	Vickrey-Clarke-Groves mechanism Nash equilibrium Markov equilibrium
Confidentiality	Limit access or place restrictions on certain types of information	Signaling game Two-player static game Repeated game	Bayesian equilibrium Nash equilibrium Subgame perfect equilibrium



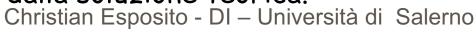


Network Security – 8/12

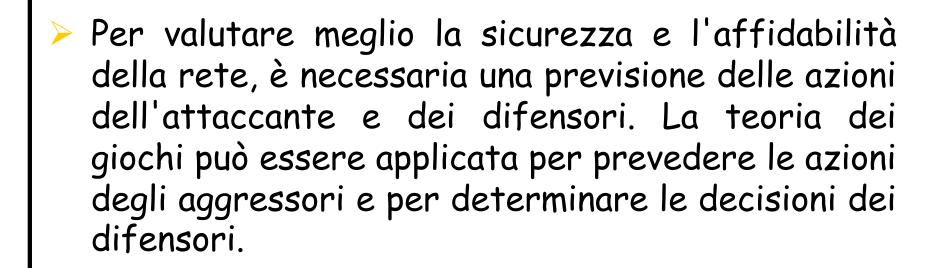
- Scenari di attacco specifici possono essere considerati nei giochi correlati alla sicurezza.
- L'interferenza nel livello di controllo di accesso medio (MAC) della rete wireless è stata anche proposta come un gioco.
 - Ciascuno dei nodi conosce solo il proprio tipo, che può essere un utente egoista o un utente malintenzionato che tenta di bloccare il canale di comunicazione, ma non il tipo di altri nodi.
- L'inceppamento è modellato come un gioco bayesiano a due giocatori in più fasi.

Network Security – 9/12

- I vantaggi delle applicazioni della teoria dei giochi per l'analisi dell'interazione attacco-difesa sono la sua semplicità e la sua facilità.
 - > Se lo scenario è semplice, l'interazione attacco-difesa può essere modellata come un gioco semplice, come un gioco statico a due giocatori o un gioco bayesiano.
- Le analisi specializzate prendono in considerazione scenari più complessi o realistici, ma sono complesse e meno solide.
 - I modelli di gioco utilizzati sono più complessi e la loro soluzione non è facile da ottenere e può richiedere una grande quantità di calcoli, e la soluzione ottenuta può deviare dalla soluzione teorica.



Network Security – 10/12



> Tale previsione viene utilizzata come input per un modulo di misurazione al fine di calcolare le metriche di sicurezza e affidabilità.





Network Security – 11/12

- > I limiti dei modelli di gioco esistenti sono:
 - 1. In genere, mancano di scalabilità. La maggior parte dei modelli di gioco per i giochi di sicurezza sono giochi a due giocatori; per gli scenari problematici con più aggressori contro più difensori, il gioco di sicurezza è nella maggior parte dei casi modellato come un gioco a due giocatori in cui tutti gli aggressori sono trattati come un giocatore, così come tutti i difensori.
 - 2. Il modello statico non è molto realistico nella maggior parte degli scenari in cui le interazioni tra gli aggressori e i difensori sono una serie di eventi;



Network Security – 12/12

- 3. I modelli stocastici presuppongono sempre che, in ogni stato, il difensore e l'attaccante possano rilevare lo stato del sistema senza errori, ma questo non è vero in molti casi realistici in cui gli IDS sono errati;
- 4. I modelli stocastici hanno delle carenze poiché presumono che gli stati del sistema siano finiti; tuttavia, gli stati del sistema sembrano essere continui sebbene alcuni modelli abbiano uno schema per suddividere lo spazio di stato continuo in parti finite;
- 5. Alcuni dei modelli di gioco stocastici non sono molto realistici perché presumono che il gioco di attacco e difesa sia un gioco a somma zero. Al contrario, un modello di gioco a somma generale è più realistico.





Security Games – 1/12

Consideriamo un semplice aeroporto con due terminal. C'è solo un'unità di polizia a proteggere i terminal e un avversario. Il Terminal 1 è più importante del Terminal 2 in questo esempio.

Adversary

Defender

	Terminal 1	Terminal 2
Terminal 1	5, –3	-1, 1
Terminal 2	- 5, 5	2, –1

La polizia può proteggere il Terminal 1 o quello 2; l'avversario può attaccare uno dei due terminal.



Security Games – 2/12

Sapendo che il Terminal 1 è più importante del Terminal 2, la polizia potrebbe scegliere di proteggere sempre il Terminal 1. Tuttavia, un avversario intelligente condurrà una sorveglianza e, dopo aver appreso che la polizia protegge sempre il Terminal 1, attaccherà il Terminal 2.

	Adversary		
		Terminal 1	Terminal 2
Derenaer	Terminal 1	5, –3	-1, 1
	Terminal 2	-5, 5	2, –1





Security Games – 3/12

Se la polizia cambiasse strategia e proteggesse sempre il Terminal 2, un avversario lo osserverebbe e successivamente attaccherebbe il Terminal 1.

1	/ lavoroally		
		Terminal 1	Terminal 2
Defender	Terminal 1	5, –3	-1, 1
	Terminal 2	- 5, 5	2, –1

Un avversario può facilmente sconfiggere qualsiasi strategia deterministica della polizia che sceglie di proteggere sempre il Terminal 1 o il Terminal 2

Security Games – 4/12

- Se la polizia randomizzasse le proprie azioni, ciò porterebbe a un risultato migliore. Un avversario che osserva la polizia non si sa esattamente dove si troverà. Ciò aumenta l'incertezza dell'avversario e migliora la ricompensa prevista per la polizia.
- Questi tipi di giochi sono chiamati giochi di Stackelberg perché la polizia si impegna innanzitutto a seguire una strategia, e un avversario agisce dopo aver effettuato la sorveglianza.

Security Games – 5/12

- Si noti che la polizia ha adottato una strategia randomizzata o "mista". L'avversario risponde con un'unica azione, non randomizzata; la reazione dell'avversario è di "pura strategia".
- Gli avversari osservano per un lungo periodo di tempo per comprendere la strategia della polizia e quindi lanciano un attacco a un obiettivo.
- L'avversario conoscerà solo la strategia generale di allocazione delle risorse a causa della precedente sorveglianza, ma non saprà esattamente come verranno allocate le risorse.



Security Games – 6/12

- Questi giochi di Stackelberg sono anche chiamati "giochi attaccante-difensore".
- Un punto chiave è che l'attaccante abbia una conoscenza perfetta della strategia mista del difensore e che l'avversario reagirà razionalmente a questa strategia, massimizzando la propria utilità attesa.
- Dai giochi di Stackelberg, passiamo ai giochi di Stackelberg bayesiani, ammettendo l'incertezza sui diversi tipi di avversari.

Security Games – 7/12

- Un tipo di avversario potrebbe considerare il Terminal 1 più importante del Terminal 2.
- Un altro tipo di avversario potrebbe considerare il Terminal 2 uguale in importanza al Terminal 1 per qualche ragione simbolica.
- Un terzo tipo di avversario potrebbe non essere in grado di attaccare efficacemente il Terminal 1, e così via.
- Non esiste solo una matrice di payoff, ma molte di esse, ciascuna corrispondente a un diverso tipo di avversario





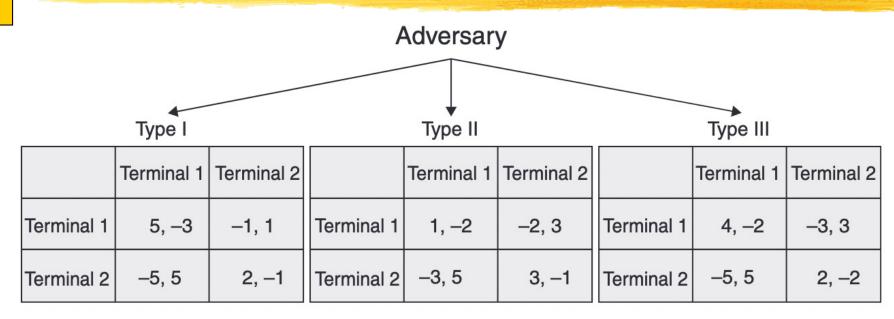
Security Games – 8/12

- Un tipo di avversario potrebbe considerare il Terminal 1 più importante del Terminal 2.
- Un altro tipo di avversario potrebbe considerare il Terminal 2 uguale in importanza al Terminal 1 per qualche ragione simbolica.
- Un terzo tipo di avversario potrebbe non essere in grado di attaccare efficacemente il Terminal 1, e così via.
- Non esiste solo una matrice di payoff, ma molte di esse, ciascuna corrispondente a un diverso tipo di avversario.





Security Games – 9/12



➤ I "giochi di sicurezza" hanno la caratteristica che ciò che è buono per l'attaccante è cattivo per il difensore e viceversa. In generale, questa tipologia di giochi non sono necessariamente a somma zero.

24

Security Games – 10/12

- L'avversario vede alcuni obiettivi come particolarmente importanti, mentre potrebbero non essere di pari importanza per la polizia.
- Un avversario potrebbe non vedere nemmeno un attacco fallito come un risultato negativo a causa della pubblicità e della paura che genera.
- L'avversario potrebbe dover sostenere un costo significativo nell'organizzare un particolare attacco che potrebbe non essere particolarmente importante per la polizia.





Security Games – 11/12

- La chiave è trovare l'allocazione ottimale delle risorse di sicurezza che ottimizzerà la ricompensa attesa del difensore. Tecnicamente, ciò che ci interessa trovare è un strong Stackelberg equilibrium (SSE).
 - In SSE, l'avversario ha una conoscenza perfetta della strategia mista del difensore e reagisce con perfetta razionalità a tale strategia, scegliendo di massimizzare la sua utilità attesa.
 - Il difensore non ha alcun incentivo a cambiare la sua strategia poiché è la strategia ottimale, e l'attaccante non ha alcun incentivo a cambiare la sua risposta perché è la risposta ottimale alla strategia mista del difensore



Security Games – 12/12

- Applicazioni di questi giochi di sicurezza si trovano in:
 - Assistant for Randomized Monitoring Over Routes (ARMOR), per la disposizione randominca dei checkpoints da parte della polizia sulle sei strade di accesso all'Aeroporto di Los Angeles;
 - Intelligent Randomization in Scheduling (IRIS), per assegnare in modo casuale gli addetti alla sorveglianza ai voli;
 - Game-theoretic Unpredictable and Randomly Deployed Security (GUARDS), per il dispiegamento randomico di unità cinofile e di controllo negli aeroporti statunitensi.





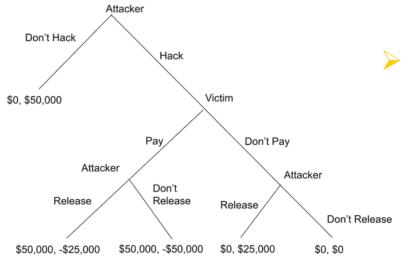
Cybersecurity Game – 1/3

- Nella sicurezza informatica, la teoria dei giochi può essere utilizzata per comprendere cosa un potenziale aggressore potrebbe tentare di fare in base ai guadagni per le sue azioni.
- In un incidente di sicurezza, il valore principale dei guadagni sono i guadagni o le perdite monetarie rispettivamente per l'aggressore e la vittima.
 - Ad esempio, un guadagno per un attacco ransomware di esempio potrebbe essere l'aggressore che guadagna \$ 50.000 mentre la vittima perde \$ 50.000 dollari più altri \$ 50.000 in tempi di inattività e indisponibilità dati.



Cybersecurity Game – 2/3

- In questa situazione abbiamo i valori assunti:
 - Prezzo del ransomware: \$ 50.000
 - Costo del tempo di inattività del sistema se l'attaccante rilascia dati: \$ 25.000
 - Costo del tempo di inattività del sistema se l'attaccante non rilascia dati: \$ 50.000



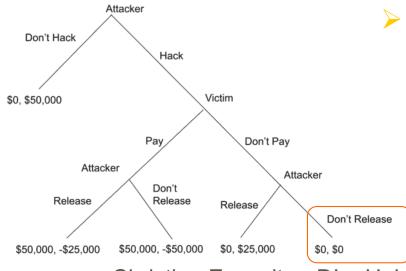
L'attaccante prende la ricompensa a sinistra, mentre la vittima prende la ricompensa a destra.





Cybersecurity Game – 2/3

- In questa situazione abbiamo i valori assunti:
 - Prezzo del ransomware: \$ 50.000
 - Costo del tempo di inattività del sistema se l'attaccante rilascia dati: \$ 25.000
 - Costo del tempo di inattività del sistema se l'attaccante non rilascia dati: \$ 50.000



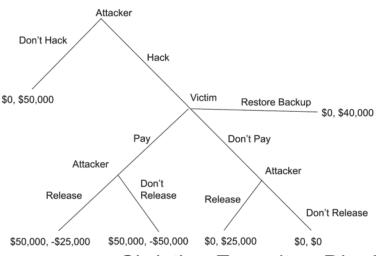
 Nel caso di equilibrio, l'attaccante muoverebbe, la vittima non pagherebbe e l'attaccante molto probabilmente non rilascerebbe i dati.





Cybersecurity Game - 3/3

- Altri fattori devono essere considerati, come la presenza di backup, la perdita maggiore può essere aggirata, con l'unica perdita rappresentata dal tempo necessario per ripristinare i backup.
- L'albero è lo stesso di cui sopra, con l'ulteriore presupposto che il costo per ripristinare il sistema utilizzando i backup sia di \$ 10.000.



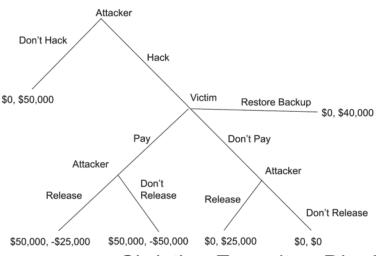
Avendo una soluzione di backup, la vittima è in grado di aggirare la seconda azione dell'attaccante.





Cybersecurity Game - 3/3

- Altri fattori devono essere considerati, come la presenza di backup, la perdita maggiore può essere aggirata, con l'unica perdita rappresentata dal tempo necessario per ripristinare i backup.
- L'albero è lo stesso di cui sopra, con l'ulteriore presupposto che il costo per ripristinare il sistema utilizzando i backup sia di \$ 10.000.



Se la vittima non paga e ripristina i backup, l'attaccante potrebbe continuare il gioco con una nuova azione, minacciando di pubblicare i dati sul dark web.



Signalling Game – 1/12

- I giochi di segnalazione sono giochi a due fasi in cui:
 - > Il Giocatore 1 (con informazioni private) muove per primo e la sua mossa viene osservata dal Giocatore 2.
 - > Il Giocatore 2 (senza conoscere le informazioni private del Giocatore 1) muove per secondo.
 - Si realizzano i payoff.
- La descrizione formale di questo gioco è:
 - Fase 0: Nature sceglie una variabile casuale t_1 , osservabile solo dal Giocatore 1, da una distribuzione $P(t_1)$.
 - Fase 1: il Giocatore 1 sceglie un'azione a_1 dall'insieme A_1 detto "messaggio"), e il Giocatore 2 osserva questa scelta di azione.
 - Fase 2: il Giocatore 2 sceglie un'azione a_2 dall'insieme A_2 .
 - Dopo la Fase 2, si realizzano i payoff: $\Pi_1(a_1, a_2; t_1)$; $\Pi_2(a_1, a_2; t_1)$.





33

Signalling Game – 2/12

Da notare che la ricompensa del Giocatore 2 dipende dal tipo di Giocatore 1.

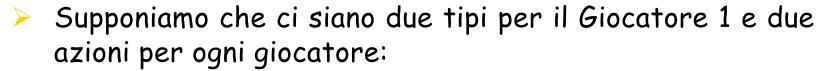
Esempi:

- \triangleright Giocatore 1: lavoratore con t_1 : capacità intrinseca, a_1 : decisione di istruzione
- Giocatore 2: azienda/e con a2: salario offerto
- Pagamenti: Π_1 = beneficio netto e Π_2 = produttività
- Giocatore 1: venditore con t_1 : vera qualità del bene, a_1 : qualità pubblicizzata
- Giocatore 2: acquirente/i con a₂: offerta offerta
- Pagamenti: Π_1 = profitto e Π_2 = beneficio netto

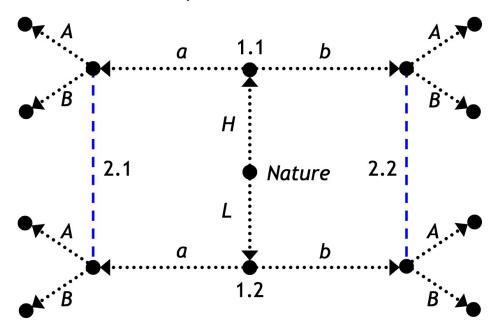




Signalling Game – 3/12



- \rightarrow $t_1 = H o t_1 = L$
- > Sia p = P(t_1 = H) con A_1 = { a, b } e A_2 = { A, B }



- Ogni giocatore ha 2
 set di informazioni e
 azioni in ognuno,
 quindi 4 strategie.
 - Un PBE è il concetto di soluzione per studiare questo tipo di gioco.





Signalling Game – 4/12



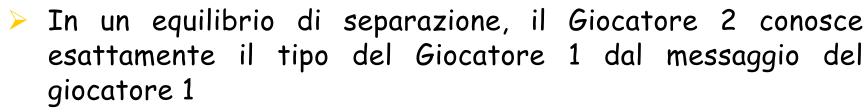
- > le convinzioni di ogni giocatore derivano da strategie usando la regola di Bayes (se possibile);
- > le strategie di ogni giocatore massimizzano il guadagno atteso date le convinzioni.
- Quando il Giocatore 1 esegue la stessa azione, indipendentemente dal suo tipo, si parla di strategia di pooling. Quando il Giocatore 1 esegue azioni diverse, a seconda del suo tipo, si parla di strategia di separazione.
- ➤ In un equilibrio di pooling, il Giocatore 2 non ottiene alcuna informazione su t₁ dal messaggio del Giocatore 1

$$\Rightarrow$$
 P₂ (t₁ = H | α_1) = P(t₁ = H) = p





Signalling Game – 5/12

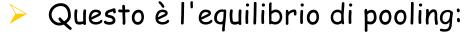


$$\Rightarrow$$
 P₂ (t₁ = H | α_1) = 0 o 1

- Supponiamo che il venditore abbia un articolo con qualità H (con probabilità p) o L (con probabilità 1 - p). Il venditore può pubblicizzare H o L.
- Supponiamo che ci siano due offerenti, che facciano sempre offerte sincere, date le loro convinzioni. Se il venditore fa sempre offerte alte. Allora: gli acquirenti non si "fideranno" mai del venditore e offriranno sempre la valutazione attesa.



Signalling Game - 6/12



$$s_1(H) = s_1(L) = H.$$

 $s_B(H) = s_B(L) = p H + (1 - p) L.$

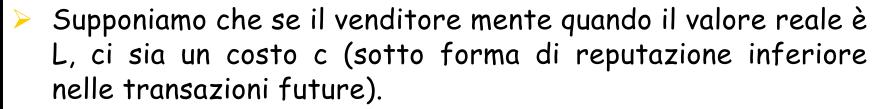
- Esiste un equilibrio in cui s_1 (t_1) = t_1 ? Ovvero che il venditore è sincero?
- In questo caso gli acquirenti fanno offerte:

$$s_B(H) = H, s_B(L) = L.$$

> se gli acquirenti usano questa strategia, il venditore preferisce pubblicizzare sempre H!



Signalling Game – 7/12



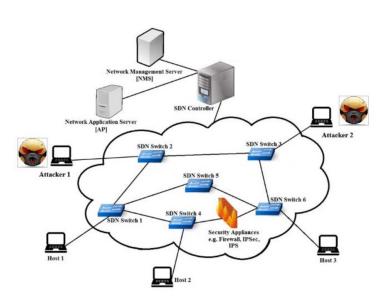
SeH-c<L,

- ➤ allora il venditore preferisce dire la verità ⇒ equilibrio di separazione.
- Questo esempio evidenzia l'importanza dei costi di segnalazione: per raggiungere un equilibrio di separazione, ci deve essere una differenza nei costi dei diversi messaggi. (Quando non ci sono costi, il messaggio risultante è chiamato "cheap talk").



Signalling Game – 8/12

Applicazione: consideriamo una rete SDN, con una serie di host e di switch SDN sotto la gestione di un controller.



- Abbiamo un gioco di segnalazione tra un target (controllore) e un mittente (legittimo/attaccante) come giocatori.
- L'attaccante (end host) ottiene l'accesso alle regole del flusso di pacchetti del controller in due modi:
- (i) durante il tempo di connessione, (ii) inviando richieste di elaborazione dei pacchetti.



Signalling Game – 9/12

- L'attaccante invia il pacchetto allo switch per l'elaborazione, quindi lo switch controlla la sua tabella di flusso per ottenere le regole di flusso.
 - > Se la tabella di flusso non contiene la struttura per quel flusso, inoltra il pacchetto al controller per generare nuove regole di flusso.
 - Al momento della generazione delle regole di flusso, l'attaccante ottiene l'accesso alla tabella di flusso del controller e ad altre informazioni importanti.
- \triangleright Il controller genera un valore di fiducia per ogni richiesta di pacchetto in arrivo e successivamente o (1) genera regole di flusso, o (2) scarta la richiesta. Il valore di fiducia di un mittente è la convinzione (θ_t) che cambia dinamicamente.





Signalling Game – 10/12

$$\theta_t = min(1, \frac{e^{(\theta_{j,0}+\theta_{j,t})/2}}{e-1})$$

- $\Theta_{j,0}$ è la convinzione statica sul mittente j al tempo t=0 (0 tentativi), e $\Theta_{j,t}$ è la convinzione dinamica sul mittente j al tempo t=t (t tentativi).
- > L'equazione indica anche che più grande è $\theta_{j,t}$; più alta è la convinzione che il mittente sia sospettoso.
- \triangleright Il modello mostra che inizialmente, il comportamento sospetto del mittente non ha un grande impatto sulla convinzione sul mittente. Tuttavia, se il comportamento sospetto continua; un piccolo cambiamento in $θ_{j,t}$ ha un impatto notevole sulla funzione di convinzione.





Signalling Game – 11/12



$$\sigma(t) = C_0 + \sum_{1 \le i \le t} C_1(i)$$

ho σ (t) è una funzione lineare in cui C_0 è il costo iniziale dell'attacco a un controllore e C_1 è il costo per unità di tempo al tentativo i. Il costo è il tempo richiesto dal mittente per effettuare un tentativo di attacco riuscito.



Signalling Game – 12/12

- Se il controller sceglie Allow, allora l'attaccante impiega banalmente meno tempo ad accedere alla regola di flusso, ovvero per un tentativo riuscito. D'altro canto, se viene scelta Ignore, il guadagno è minore, il che spinge l'attaccante a sospendere la sua attività per un po' di tempo.
- \triangleright Il controller deve agire con attenzione in modo che né il mittente legittimo subisca danni né l'attaccante esegua attività dannose. L'azione del controller comporta un costo definito dal parametro $ψ_t$ e rappresentato come segue:

$$\psi_t = \sum_{1 \le i \le t} \rho_i A_i = \psi_{t-1} + \rho_t A_t$$

 $ho_i A_i$ indica il costo del controllore per intraprendere l'azione A_i al tentativo i.

