

Estudio de los Procesos de Decisión Markoviana para los Sistemas de Control en Red

Study of Markovian Decision Processes for Network Control Systems

Carlos Pillajo A.

Universidad Politécnica Salesiana

cpillajo@ups.edu.ec

Roberto Hincapié

Universidad Pontificia Bolivariana

roberto.hincapie@upb.edu.co

Resumen

Actualmente estamos migrando de un mundo digital a un mundo cibernético en donde muchas acciones de monitoreo y control se están realizando a través de redes compartidas como el internet, de ahí surgen los sistemas de control en red para lo cual se necesita saber cómo implementar procesos de decisión Markoviana en este tipo de sistemas, por lo tanto, vamos a realizar un estudio de cómo aplicar la teoría de los Procesos de Decisión Markoviana en los Sistemas de Control en Red, esperamos con este estudio dar un aporte de actualización de la teoría de control para procesos sujetos a restricción de comunicación entre el sensor, el controlador y el actuador, los que no se encuentran físicamente en un mismo sitio y se comunican mediante canales de comunicación compartida.

Los Sistemas de control en red (NCS) han sido una de las principales líneas de investigación donde se centra el mundo académico, así como en aplicaciones industriales. Los NCS ha tomado la forma de un área multidisciplinar. Presentamos las diferentes formas de Procesos de Decisión de Markov con aplicaciones en los NCS, desde los recientes resultados en este campo, así como los sistemas de NCS basado en modelo, realizando una revisión de la estabilidad en los NCS

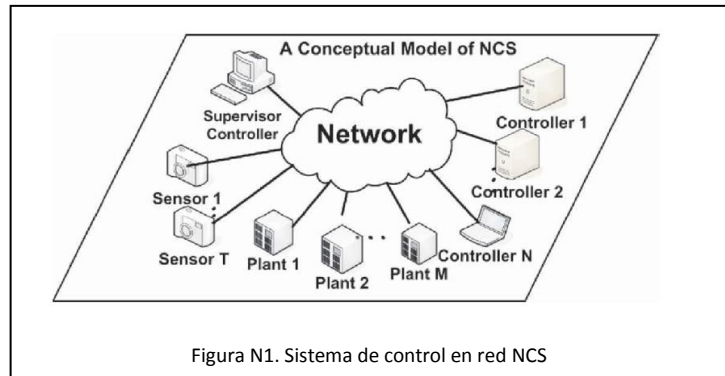
Palabras claves: Sistemas de Control en Red, Procesos de Decisión Markoviana.

INTRODUCCION

Un sistema de control es un dispositivo o conjunto de dispositivos para administrar, comandar ó regular el comportamiento de otros dispositivos a través de canales de comunicación compartido. Durante muchos años, los investigadores nos han dado las estrategias precisas y óptimas de control que

salen de la teoría de control clásica, a partir del control de lazo abierto a las estrategias de control sofisticados basados en algoritmos genéticos.

El advenimiento de las redes de comunicación, sin embargo, introdujo el concepto de controlar de forma remota un sistema, que dio a luz a los sistemas de control en red (NCS). La definición clásica de NCS puede ser como sigue: Cuando un sistema de control de realimentación tradicional



está cerrado a través de un canal de comunicación, que puede ser compartida con otros nodos fuera del sistema de control, entonces el sistema de control se llama un NCS (Liu, 2008). Un NCS también se puede definir como un sistema de control de retroalimentación en el que los bucles de control están cerrados a través de una red en tiempo real. La característica definitoria de un NCS es que la información (entrada de referencia, producción de la planta, entrada de control, etc.) se intercambian utilizando una red entre los componentes del sistema de control (Figura N1).

Ventajas y aplicaciones de control sobre la Red

Desde hace muchos años, las tecnologías de redes de datos se han aplicado ampliamente en aplicaciones de control industrial y militar. Estas aplicaciones incluyen plantas de fabricación, automóviles y aviones. Conexión de los componentes del sistema de control en estas aplicaciones, tales como sensores, controladores, y actuadores, a través de una red puede reducir eficazmente la complejidad de los sistemas, con inversiones económicas nominales. Además, los controladores de red permiten que los datos sean compartidos de manera eficiente, eliminan cableado innecesario. Es fácil agregar más sensores, actuadores y controladores con muy poco costo y sin cambios estructurales pesados a todo el sistema. Lo más importante es que se conectan al ciberespacio para la ejecución de tareas en un espacio físico desde una distancia de fácil acceso (una forma de tele-presencia). Estos sistemas son cada vez más realizables hoy y tienen una gran cantidad de posibles aplicaciones, incluyendo exploraciones espaciales, la exploración terrestre, la automatización de fábrica, el diagnóstico remoto y resolución de problemas, entornos peligrosos, las instalaciones experimentales, los robots domésticos, automóviles, aviones, monitoreo de la planta de fabricación, hogares de ancianos u hospitales, tele-robótica y tele-operación, sólo para nombrar unos pocos.

PROCESOS DE DECISION DE MARKOV EN LOS SISTEMAS DE CONTROL EN RED (NCS)

Los NCS funcionan como sistemas estocásticos debido a la aleatoriedad en los ambientes controlados. Para el tiempo de servicio largo y bajo costo de mantenimiento, NCS requieren métodos adaptativos y sólidos para abordar el intercambio de datos. Este estudio examina numerosas aplicaciones de la plataforma de procesos de decisión de Markov (MDP), una herramienta de toma de decisiones de gran alcance para el desarrollo de algoritmos adaptativos. (Abu Alsheikh, Hoang, Niyato, Tan, & Lin, 2015)

El modelado de MDP ofrece los siguientes beneficios generales para operaciones con NCS:

1) Los Sistemas Controlados en Red (NCS's) son sistemas que deben optimizar los recursos, por lo que no es posible que un sensor en una planta este emitiendo datos al controlador remoto si no existe cambio en la planta, eso provoca consumo de ancho de banda innecesario. Por lo tanto, el uso de PDM para optimizar dinámicamente las operaciones de red para adaptarse a las condiciones físicas resultados en la utilización de recursos mejorando significativamente.

2) El modelo MDP permite un diseño equilibrado de diferentes objetivos, por ejemplo, minimizando el consumo de energía y maximizar el uso del canal de comunicación sujeto a restricciones, limitaciones en cuanto a su uso, pérdida de paquetes y retardos

3) Nuevas aplicaciones de NCS interactúan con entidades móviles que aumentan significativamente la dinámica del sistema. Por ejemplo, utilizando una pasarela móvil de recogida de datos introduce muchos retos de diseño. En este caso, el método MDP puede explorar la correlación temporal de los objetos en movimiento y predecir sus posiciones futuras.

4) La solución de un modelo de MDP, referido como una política, se puede implementar en base a una tabla de consulta. Esta tabla se puede almacenar en la memoria del nodo sensor para realizar operaciones en línea con mínima complejidad. Por lo tanto, el modelo MDP se puede aplicar incluso para nodos diminutos y limitados de recursos sin ningún requisito de alta computación. Por otra parte, las soluciones casi óptimas se pueden derivar para aproximar las políticas óptimas de decisión que permite el diseño de algoritmos para NCS con menos cargas de cálculo.

5) Los Procesos de Decisión Markoviana (MDP) son flexibles con muchas variantes que pueden caber las condiciones distintas en las aplicaciones NCS. Por ejemplo, nodos sensores generalmente producen lecturas ruidosas, por lo tanto, dificulta el proceso de toma de decisiones. Con estas observaciones imprecisas, una de las variantes del MDP, es decir, proceso de decisión de

Markov parcialmente observable (POMDP), se puede aplicar para alcanzar la mejor política operacional.

Este trabajo se centra en las aplicaciones de MDP en NCS. La principal diferencia entre un MDP y una cadena de Markov es que la cadena de Markov no considera las acciones y recompensas. Por lo tanto, sólo se utiliza para análisis de rendimiento. Por el contrario, el MDP se utiliza para la optimización estocástica, es decir, para obtener las mejores acciones a tomar dada objetivos particulares y, posiblemente un conjunto de restricciones. (Abu Alsheikh, Hoang, Niyato, Tan, & Lin, 2015)

Procesos de Decisión Markoviana

Un proceso de decisión de Markov (MDP) es un modelo de optimización para la toma de decisiones bajo incertidumbre (Inria, 2004). El MDP describe un proceso de decisión estocástico de un agente de interacción con un entorno o sistema. En cada momento de la decisión, el sistema se mantiene en un cierto estado s y el agente elige una acción a que está disponible en este estado. Después se realiza la acción, el agente recibe una R recompensa inmediata y el sistema transita a un nuevo estado s_0 estado de acuerdo con la probabilidad de transición $P_{a s y; s_0}$. Por ejemplo, el MDP puede optimizar la decisión de envíos de paquetes de control en un NCS.

A. El proceso de Decisión Markoviana El Marco

El MDP se define por una tupla $[S; A; P; R; T]$ donde,

- S es un conjunto finito de estados,
- A es un conjunto finito de acciones,
- P es una función de probabilidad de transición de estado s para el estado S_0 después de que se tome la acción a ,
- R es la recompensa que se obtiene inmediatamente después de la acción de una se hace,
- T es el conjunto de la decisión, que puede ser finito o infinito.

π denota una "política", que es un mapeo de un estado a una acción. El objetivo de un MDP es encontrar una política óptima para maximizar o minimizar una cierta función objetivo. Un MDP puede ser horizonte temporal finito o infinito. Para el tiempo finito horizonte MDP, una política óptima π^* para maximizar la recompensa total esperado se define como sigue:

$$\max V_{\pi}(s) = \mathbb{E}_{\pi, s} \left[\sum_{t=1}^T \mathcal{R}(\acute{s}_t | s_t, \pi(a_t)) \right] \quad (1)$$

donde s_t y a_t son del estado y la acción en el tiempo t , respectivamente.

Para el tiempo infinito horizonte MDP, el objetivo puede ser para maximizar la recompensa total descontada o para maximizar la recompensa promedio. Lo anterior se define como sigue:

$$\max V_{\pi}(s) = \mathbb{E}_{\pi,s} \left[\sum_{t=1}^T \gamma^t \mathcal{R}(\acute{s}_t | s_t, \pi(a_t)) \right] \quad (2)$$

Mientras que el segundo se expresa como sigue:

$$\max V_{\pi}(s) = \liminf_{T \rightarrow \infty} \frac{1}{T} \mathbb{E}_{\pi,s} \left[\sum_{t=1}^T \gamma^t \mathcal{R}(\acute{s}_t | s_t, \pi(a_t)) \right] \quad (3)$$

En este caso, γ es el factor de descuento y $\mathbb{E}[\cdot]$ Es la función de las expectativas.

B. Las soluciones de PDM

Aquí presentamos los métodos de solución de MDPs con recompensa total descontada. Los algoritmos para MDPs con la recompensa promedio se pueden encontrar en (Inria, 2004).

1) Soluciones para Procesos de Decisión de Markov de Horizonte finito

temporal: En un tiempo finito horizonte MDP, el funcionamiento del sistema se lleva a cabo en un período de tiempo conocido. En particular, el sistema comienza a s_0 estado y sigue funcionando en los próximos períodos T . La política óptima π^* es maximizar $V_{\pi}(s)$ en (1). Si denotamos $v^*(s)$ como la recompensa máxima alcanzable en el estado s , entonces podemos encontrar $v^*(s)$ en cada estado de forma recursiva mediante la resolución de ecuaciones óptimas de Bellman:

$$v_t^*(s) = \max_{a \in A} \left[\mathcal{R}_t(s, a) + \sum_{s' \in S} \mathcal{P}_t(s' | s, a) v_{t+1}^*(s') \right] \quad (4)$$

Sobre la base de las ecuaciones óptimas de Bellman, dos enfoques típicos para tiempo de horizonte finito de MDPs existen.

Inducción hacia atrás: También conocido como un enfoque de programación dinámica, es el método más popular y eficaz para la solución de las ecuaciones de Bellman. Dado que el proceso se detuvo en un período conocido, lo primero que puede determinar la acción óptima y la función de valor óptimo en el último período de tiempo. A continuación, de forma recursiva obtenemos las acciones óptimas para períodos anteriores se remontan al primer periodo basado en las ecuaciones óptimas de Bellman.

Inducción hacia adelante: Este método de inducción hacia adelante también se conoce como un enfoque de valor iteración. La idea es dividir el problema de

optimización basado en el número de pasos para ir. En particular, dada una política óptima para t-1 pasos de tiempo para ir, calculamos los valores de Q para k pasos para ir. Después de eso, se puede obtener la política óptima basada en las siguientes ecuaciones:

$$Q_t(s, a) = R(s, a, s') + \sum_{s'} P(s, a, s') v_{t-1}^* | (s'),$$

$$v_t^*(s) = \max_{a \in A} Q_t^*(s, a) \text{ and } \pi_t^*(s) = \operatorname{argmax}_{a \in A} Q_t^*(s, a),$$

donde $v_t(s)$ es el valor de estado s y $Q_t(s; a)$ es el valor de actuar en un estado s . Se llevará a cabo este proceso hasta que se alcance el último período.

Ambos enfoques tienen la misma complejidad que depende del horizonte de tiempo de un MDP. Sin embargo, se utilizan de manera diferente. Inducción hacia atrás es especialmente útil cuando se sabe el estado de PDM en el último período. Por el contrario, la inducción hacia adelante se aplica cuando sólo sabemos el estado inicial.

2) Las soluciones para los procesos de Decisión Markoviana para horizonte de tiempo infinito: Solución de un infinito horizonte de tiempo MDP es más compleja que la de un tiempo finito horizonte MDP. No obstante, el infinito horizonte de tiempo MDP es el más utilizado porque en la práctica el tiempo de funcionamiento de los sistemas a menudo se desconoce y asume que es infinita. Se propusieron muchos métodos de solución.

Valor iteración (VI): Este es el método más eficiente y ampliamente utilizado para resolver un tiempo infinito horizonte descuento MDP. Este método tiene muchas ventajas, por ejemplo, la convergencia rápida, facilidad de implementación, y es especialmente una herramienta muy útil cuando el espacio de estados del PDM es muy grande. Al igual que el método de inducción hacia adelante de un tiempo finito horizonte MDP, este enfoque también fue desarrollado en base a la programación dinámica. Sin embargo, para infinito horizonte de tiempo MDP, ya que el horizonte de tiempo es infinito, en lugar de ejecutar el algoritmo para todo el horizonte de tiempo, tenemos que utilizar un criterio de parada (por ejemplo, $\|v_t^* - v_{t-1}^*\| < \varepsilon$) para garantizar la convergencia.

Política de iteración (PI): La idea principal de este método es generar una secuencia de mejora de las políticas. Se inicia con una política arbitraria y actualiza la política hasta que converge. Este enfoque consiste en dos pasos principales, a saber, la evaluación de políticas y la mejora de las políticas. En primer lugar, resolver las ecuaciones lineales para encontrar la recompensa descontado esperado bajo la política π y luego elegir la política de decisiones mejora para cada estado. En comparación con el método de iteración de valor,

este método puede tener un menor número de iteraciones para converger. Sin embargo, cada iteración toma más tiempo que el del método de valor de repetición porque el método de iteración de política requiere la resolución de ecuaciones lineales.

La programación lineal (LP): A diferencia de los métodos anteriores, el método de programación lineal tiene como objetivo encontrar una política estática a través de la solución de un problema de programación lineal. Después de que el programa lineal se resuelve, se puede obtener el valor óptimo $v^*(S)$, a partir de la cual se puede determinar la política óptima $\pi^*(s)$ en cada estado. El método de programación lineal es relativamente ineficiente en comparación con los métodos de iteración de valor y de política cuando el espacio de estados es grande. Sin embargo, el método de programación lineal es útil para MDPs con limitaciones ya que las restricciones pueden ser incluidos como ecuaciones lineales en el programa lineal (Inria, 2004).

Método de aproximación: programa dinámico aproximado fue desarrollado para grandes MDPs. El método aproxima las funciones de valor (si las funciones de política o funciones de valor) suponiendo que estas funciones pueden ser caracterizados por un número razonable de parámetros. Por lo tanto, podemos buscar los valores de los parámetros óptimos para obtener la mejor aproximación, por ejemplo, como se da en (Sutton, 1992).

El aprendizaje en línea: Los métodos antes mencionados se llevan a cabo de una manera fuera de línea (es decir, cuando se proporciona la función de probabilidad de transición). Sin embargo, no pueden utilizar si no se conoce la información de dichas funciones. Se propusieron algoritmos de aprendizaje para abordar este problema (Sigaud & Buffet, 2013). La idea se basa en el método basado en la simulación que evalúa la interacción entre un agente y sistema. Luego, el agente puede ajustar su comportamiento para lograr su objetivo (por ejemplo, el ensayo y error).

C. Desafíos de la aplicación PDM para NCS

Los NCSs encuentran cada vez más y nuevas aplicaciones que sirve como una plataforma clave en muchas tecnologías inteligentes y el Internet de las Cosas (IoT). Esto introduce continuamente desafíos de diseño abierto en el que los PDM se puede utilizar para tomar decisiones. En esta sección, se discuten algunos problemas de investigación abiertas que no se han estudiado detalladamente en la bibliografía, y que requieren mayor atención de la investigación.

El marco MDP es una poderosa herramienta de análisis para abordar los problemas de optimización estocástica. El marco MDP ha demostrado su aplicabilidad en muchas implementaciones del mundo real, tales como finanzas, agricultura, deportes, etc (Hespanha, Naghshtabrizi, & Xu, 2007). Sin

embargo, todavía hay algunas limitaciones que requieren mayor estudio de investigación.

1) La sincronización de tiempo: La mayoría de los estudios existentes asumen sincronización de tiempo perfecta entre los nodos. Este supuesto permite a los nodos de la red para construir un ciclo MDP unificado (sentido actual estado, tomar decisiones y realizar acciones, percibir nuevo estado, etc). Los mecanismos para abordar estas cuestiones deben ser desarrollados.

2) El Curso de dimensionalidad: Este es un problema inherente de los PDM cuando el espacio de estado y / o el espacio de acción se hacen grandes. En consecuencia, no podemos resolver PDM directamente mediante la aplicación de métodos de solución estándar.

3) Estacionariedad y Modelos variable en el tiempo: Se supone que las probabilidades de transición del MDP y la función de la recompensa son tiempo invariable. Sin embargo, en algunos sistemas, esta suposición puede ser inviable. Hay dos métodos generales para hacer frente a las probabilidades de transición no estacionarias en problemas de decisión de Markov. En la primera solución, un algoritmo de aprendizaje en línea, por ejemplo se utiliza para actualizar las probabilidades de transición del estado y la función de recompensa basado en los cambios en el entorno.

Por lo tanto, la literatura existente carece de visión de experimentos a largo funcionamiento el uso de bancos de pruebas del mundo real e implementaciones para evaluar el rendimiento de todo el sistema en condiciones cambiantes, y es preciso, por tanto, bancos de pruebas rigurosas para las validaciones sistemáticas de los modelos de MDP.

Redes de Procesos de Decisión Markoviana con retrasos

Consideramos un sistema de control de red, donde cada subsistema evoluciona como un proceso de decisión de Markov con algunas entradas adicionales de otros sistemas. Cada subsistema está acoplado a sus vecinos a través de enlaces de comunicación sobre las que las señales se retrasaron, pero se transmiten de otra manera libre de ruidos. Un controlador centralizado recibe retrasó la información de estado de cada subsistema. La acción de control se aplica a cada subsistema tiene efecto después de un cierto retraso y no inmediatamente.

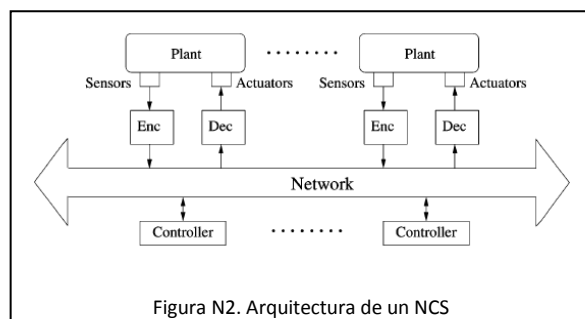
Estamos interesados en el control de una red interconectada de subsistemas. Cada subsistema se modela como un proceso de decisión de Markov (MDP), y el sistema en general se conoce como un proceso de toma de red de Markov, que se utiliza para modelar una variedad de problemas de control. Esta nota técnica muestra que para MDPs en red, el controlador óptimo

es una función de un número finito de observaciones pasadas. Se demuestra que para MDPs conectados en red, los resultados dependen sólo de la estructura de red y los retardos asociados. En el marco de esta nota técnica (Adlakha, Lall, & Goldsmith, 2012), los subsistemas están acoplados entre sí a través de enlaces de comunicación que son libres de ruido, líneas de retardo puro, sin pérdidas de paquetes u observaciones ruidosas. Los retardos son fijos, pero puede ser diferente para cada interconexión. Suponemos que cada subsistema tiene un espacio de estados finitos. Un controlador centralizado recibe retrasó mediciones del estado de cada subsistema y calcula una acción de control óptima para ser aplicado a cada subsistema, que tiene efecto después de un cierto retraso. Aunque el controlador recibe información del estado de cada subsistema, cada uno de estos estados se retrasa por diferentes cantidades, por lo que el estado actual de cada subsistema no está disponible para el controlador. Este sistema por lo tanto se puede representar como un proceso parcialmente observado de decisión de Markov (POMDP).

Diseño de control óptimo para MDPs y POMDPs se ha estudiado ampliamente en la literatura (Monahan, 1982). Hay dos enfoques estándar de un control óptimo de POMDPs. El primer enfoque genera una política que es una función de toda la historia de observación; esta historia se llama un estado de información y crece sin límite a medida que aumenta el tiempo. En el segundo enfoque, el controlador es una función de la creencia de estado que es la distribución posterior del estado actual del sistema de acondicionado en toda la historia de observación.

RESULTADOS EN LOS SISTEMAS DE CONTROL EN RED

Sistemas de control en red (NCS) son sistemas para que la comunicación entre los sensores, actuadores y controladores con el apoyo de una comunicación digital de banda limitada compartida, como se muestra en la Figura N2. Revisamos varios resultados recientes sobre la estimación, análisis y síntesis de controlador para NCS.



En consecuencia, NCSs han ido encontrando aplicación en una amplia gama de áreas, tales como las redes móviles de sensores, la cirugía a distancia, hápticos colaboración a través de Internet, etc. Sin embargo, el uso de un

contraste network - in compartida a la utilización de varias conexiones independientes dedicados introduce nuevos retos, del control sobre las redes como una de las futuras orientaciones clave para el control.

NCSs encuentran en la intersección de las teorías de control y comunicación. Tradicionalmente, la teoría de control se centra en el estudio de los sistemas dinámicos interconectados vinculados a través de canales ideal, mientras que la teoría de comunicación estudia la transmisión de información a través de canales imperfectos. Se necesita una combinación de estos dos marcos para modelar NCSs. Este estudio está escrito principalmente desde la perspectiva de los controles y los intentos de manera sistemática abordar varios aspectos clave que hacen NCSs distinto de otros sistemas de control.

a) Canales de Banda Limitada .- Cualquier red de comunicación sólo puede llevar a una cantidad limitada de información por unidad de tiempo. En muchas aplicaciones, esta limitación plantea limitaciones significativas en el funcionamiento de NCS. La mayoría de los resultados analizados en este estudio están motivados por la observación de que, en la mayoría de las redes digitales, los datos se transmiten en unidades llamadas paquetes y el envío de un solo bit o de varios cientos de bits consume la misma cantidad de recursos de la red. Esta observación conduce a un punto de vista alternativo de los canales de banda limitada, en el que un canal puede transmitir un número finito de paquetes por unidad de tiempo (tasa de paquetes), pero cada paquete puede llevar un gran número de bits (posiblemente infinitamente muchos). Aunque estos canales tienen capacidad infinita, la estabilidad del sistema en lazo cerrado y el rendimiento es todavía limitada por la red.

b) Toma de muestras y de retardo.- para transmitir una señal continua en el tiempo través de una red, la señal debe ser muestreada, codificado en un formato digital, transmitida por la red, y finalmente los datos debe ser decodificado en el lado receptor. Este proceso es significativamente diferente de la toma de muestras periódica habitual en el control digital. El retardo global entre el muestreo y decodificación eventual en el receptor puede ser muy variable porque tanto los retardos de acceso de red (es decir, el tiempo que toma para una red compartida para aceptar datos) y los retardos de transmisión (es decir, el tiempo durante el cual los datos están en de tránsito dentro de la red) dependen de las condiciones de red altamente variables, tales como la congestión y la calidad del canal. En algunos NCS, los datos transmitidos son marca de tiempo, lo que significa que el receptor puede tener una estimación de la duración de la demora y adoptará las medidas correctoras pertinentes. Un

número significativo de los resultados han tratado de caracterizar un máximo límite superior del intervalo de muestreo para los que se puede garantizar la estabilidad. Estos resultados implícitamente intentan minimizar la tasa de paquetes que es necesaria para estabilizar un sistema a través de la retroalimentación.

c) Pérdida de Paquete.- Otra diferencia significativa entre NCS y control digital estándar es la posibilidad de que los datos se pueden perder durante el tránsito a través de la red. Típicamente, los abandonos de paquetes son el resultado de errores de transmisión de los enlaces de red físicos (que es mucho más común en inalámbrico que en las redes de cable) o de desbordamientos de búfer debido a la congestión. Retardos de transmisión largas a veces resultan en la reordenación de paquetes, lo que esencialmente equivale a un desertor de paquetes si el receptor descarta las llegadas. Protocolos de transmisión fiables, como TCP, garantizan la eventual entrega de paquetes. Sin embargo, estos protocolos no son apropiados para NCSs ya que la retransmisión de datos antiguos generalmente no es muy útil.

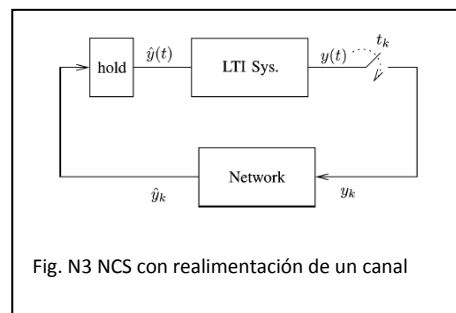
A continuación revisaremos la estabilidad de los sistemas realimentados, este problema se da por los escenarios en los que sensores, controladores y actuadores son colocados y utilizados en redes de comunicación compartida.

Muestreo y retardo.

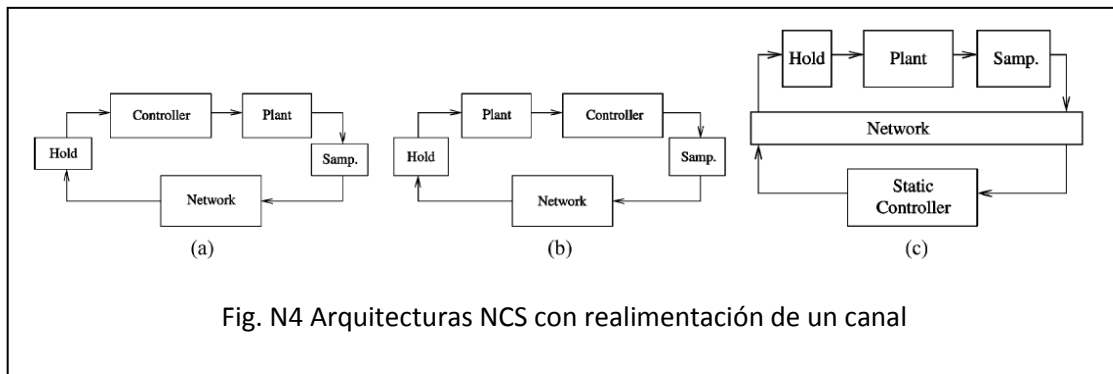
La retroalimentación de un canal en un NCS figura N3 ha sido comumente utilizada para investigar los efectos del muestreo y los retardos en la estabilidad de los NCS. Los sistemas LTI, linealmente invariantes en el tiempo modela el par planta/controlador por la siguiente sistema de tiempo continuo:

$$\dot{x} = Ax + B\hat{y}, \quad y = Cx$$

Esta retroalimentación de un canal NCS puede capturar varias configuraciones NCS. La señal puede ser considerado como un vector de mediciones del sensor y \hat{y} como la entrada a un controlador de tiempo continuo colocado con el actuador, como en la Fig. 4 (a) (Montestruque & Antsaklis, 2003).



Alternativamente, \hat{y} puede ser visto como la entrada a los actuadores e y como la señal de control deseada calculada por un controlador colocada con los sensores, como en la Fig. 4 (b). En cualquier caso, x incluiría los estados de la planta y el controlador. El diagrama de bloques en la Fig. 4 también captura el caso de un controlador estático que no está en el mismo emplazamiento de los sensores ni con los actuadores como en la Fig. 4 (c), porque un controlador sin memoria podría ser movido junto a los actuadores, sin afectar a la estabilidad del bucle cerrado (Montestruque & Antsaklis, 2004).



En la retroalimentación de un canal NCS en la Fig. 3, la señal $y(t)$ se muestrea al tiempo $\{t_k: k \in \mathbb{N}\}$ las muestras $y_k := y(t_k), \forall k \in \mathbb{N}$ se envían a través de la red. En una red sin pérdidas, tenemos

$$\hat{y}_k = y_k, \quad \forall k \in \mathbb{N}$$

Pero las muestras solamente llegan al destino después de un (posiblemente variable) retardo de $\tau_k \geq 0$. a este tiempo $\hat{y}(t)$ es actualizado, dando lugar a

$$\hat{y}(t) = \begin{cases} \hat{y}^{k-1}, & t \in [t_k, t_k + \tau_k) \\ \hat{y}_k, & t \in [t_k + \tau_k, t_{k+1}) \end{cases}$$

Donde asumimos que los retardos de la red son siempre más pequeños que un intervalo de la muestra. $t_k + \tau_k < t_{k+1}, \forall k \in \mathbb{N}$

Cuando la planta de lazo abierto es inestable, generalmente es inestable si el intervalo entre los tiempos de muestreo se hace muy grande. En vista de esto, el trabajo significativo se ha dedicado a la búsqueda de límites superiores en $t_{k+1} - t_k, \forall k \in \mathbb{N}$ para el cual se puede garantizar la estabilidad. Estos límites superiores se llaman a veces el intervalo de transferencia máxima permitida (MATI) (Walsh, Ye, & Bushnell, 2002).

Retrasos de más de un intervalo de muestreo pueden dar lugar a más de un \hat{y}_k (o ninguno) que llegan durante un intervalo de muestreo simple, por lo que la utilización de fórmulas recursivas se dificulta. Todos los resultados analizados en esta sección se basan en una reducción de los NCS a algún tipo de sistema de tiempo discreto como ; para simplificar, vamos a suponer, por tanto, implícitamente retrasos menores de un intervalo de muestreo.

Perdida de Paquetes

Abandonos de paquetes se pueden modelar ya sea como estocástico o fenómenos deterministas. El modelo estocástico simple asume que los abandonos son realizaciones de un proceso de Bernoulli. Cadenas de Markov de estado finito se pueden utilizar para modelar abandonos correlacionados (Smith & Seiler, 2003) y procesos de Poisson pueden ser usados para modelar abandonos estocásticos en tiempo continuo (Xu & Hespanha, 2005). También se han propuesto modelos deterministas de los abandonos, ya sea especificada en términos de promedios de tiempo (Zhang, Branicky, & Phillips, 2001) o en términos del peor de los casos límites en el número de abandonos consecutivos.

Consideremos nuevamente la realimentación de un canal NCS en la figura N3, con un par planta / controlador, para lo cual la señal y se muestrea al tiempo $\{t_k: k \in \mathbb{N}\}$ y las muestras $\{y_k := y(t_k)\}$ se envían a través de la red. Cuando se dejan caer los paquetes, el modelo de red en debe ser cambiado. A menudo se asume que cuando se deja caer el paquete que contiene el $\{y_k\}$, el NCS utiliza el valor anterior de $\{\hat{y}_k\}$. Esto corresponde a sustituir el modelo de red a uno con pérdidas por:

$$\hat{y}_k = \theta_k y_k + (1 - \theta_k) \hat{y}_{k-1} = \begin{cases} y_k & \theta_k = 1 \text{ (sin paquetes perdidos)} \\ y_k & \theta_k = 0 \text{ (con paquetes perdidos)} \end{cases} \forall k \in \mathbb{N}$$

Donde $\theta_k = 0$ cuando hay un paquete perdido al tiempo k y $\theta_k = 1$ en otro caso. Hadjicostis (Hadjicostis & Touri, 2002) asume en cambio \hat{y}_k se pone a cero cuando se deja caer el paquete que contienen y_k , es decir, $\hat{y}_k = \theta_k y_k, \forall k \in \mathbb{N}$.

CONCLUSIÓN Y DIRECCIÓN DE FUTURAS INVESTIGACIONES

Se ha presentado una colección de resultados para determinar la estabilidad de bucle cerrado de NCS en presencia de muestreo red, retrasos y abandonos de

paquetes. Consideramos una variedad de hipótesis sobre los efectos de la planta y de la red, lo que llevó a los modelos de circuito cerrado representados por Modelos Markovianos, sistemas de variables en el tiempo lineal, los sistemas no lineales con reinicios, los sistemas dinámicos asíncronos (ADSs) de conmutación, sistemas lineales invariantes en el tiempo con estocástico estructurado incertidumbre, y los sistemas lineales con entradas retardados. Muchos de los resultados presentados se basan en técnicas Lyapunov y sólo proporcionan condiciones suficientes para la estabilidad de la NCS

En este estudio no abordó algunas cuestiones importantes en NCS, tales como velocidad de bits y cuantificación. El problema de determinar la tasa de bits mínima necesaria para la estabilización ha sido resuelto exactamente para las plantas lineales, pero sólo los resultados conservadoras se han obtenido para las plantas no lineales. Cuantificación se vuelve especialmente importante para las redes diseñados para transportar paquetes muy pequeños con poca sobrecarga, ya que para este tipo de redes se puede ahorrar ancho de banda mediante la codificación de mediciones o señales de actuación con un pequeño número de bits.

NCS han estado atrayendo un interés significativo en los últimos años y seguirá haciéndolo durante los próximos años. Con la llegada de los procesadores barato, pequeño y de bajo consumo con capacidades de comunicación, se ha hecho posible dotar de sensores y actuadores con potencia de procesamiento y la capacidad de comunicarse con los controladores quitan a través de redes de usos múltiples. En vista de esto, conjeturamos que en un futuro próximo NCS se convertirá en la norma, en sustitución de los actuales sistemas de control de tasa fija digitales que dependen de conexiones dedicadas entre sensores, controladores y actuadores. Resultados aún faltan para superar algunos de los desafíos planteados por NCS. Entre ellos destacamos los siguientes. Ha sido un trabajo significativo en NCS con tasa de muestreo variable, pero la mayoría de los resultados investigar la estabilidad para el peor caso determinado intervalo entre los tiempos de muestreo consecutivos. Esto generalmente conduce a resultados conservadores que podrían mejorarse teniendo en cuenta una caracterización estocástica para los tiempos intersampling. La mayoría del trabajo se ha dedicado a la determinación de la estabilidad de NCS, mientras que las cuestiones relacionadas con el rendimiento se han descuidado un poco. El diseño de controladores para NCS también ha pasado por alto, ya que muchos investigadores comienzan con un controlador que ha sido diseñado, ignorando los desafíos introducidos por NCS y luego investigar

en qué medida dichos controladores pueden garantizar la estabilidad a pesar de la red

Después de tener una visión general de las diferentes categorías, componentes y aplicaciones de NCS, que ahora describimos los diferentes retos y cuestiones a tener en cuenta para una NCS fiable. Podemos clasificar ampliamente aplicaciones NCS en dos categorías como: (1) aplicaciones sensibles al tiempo o control de tiempo crítico como las operaciones militares, espaciales y de navegación; (2) en tiempo insensible o control en tiempo no real, como el almacenamiento de datos, la recopilación de datos del sensor, e-mail, etc. Sin embargo, la fiabilidad de la red es un factor importante para ambos tipos de sistemas. La red puede introducir niveles poco fiables y que dependen del tiempo de servicio en términos de, por ejemplo, demoras, jitter, o pérdidas. Calidad de servicio (QoS) puede mejorar el comportamiento de la red en tiempo real, pero el comportamiento de la red sigue siendo objeto de interferencia (sobre todo en los medios de comunicación sin cable), a los transitorios de enrutamiento, y para los flujos agresivos. A su vez, los caprichos de la red pueden poner en peligro la estabilidad, la seguridad y el rendimiento de las unidades en un entorno físico. Un problema difícil en el control de los sistemas basados en la red son los efectos de retardo de red. El tiempo para leer una medición del sensor y para enviar una señal de control a un actuador a través de la red depende de las características de la red, como los planes de topología y enrutamiento. La gravedad del problema de retardo se agrava cuando se produce la pérdida de datos durante una transmisión. Por otra parte, los retrasos no sólo se degradan el rendimiento de un sistema de control basado en la red, pero también pueden desestabilizar el sistema.

En estos problemas, los nodos sensores son para tomar decisiones optimizados de un conjunto de estrategias accesibles para lograr objetivos de diseño. Este estudio examina numerosas aplicaciones de la plataforma de procesos de decisión de Markov (MDP), una herramienta de toma de decisiones de gran alcance para el desarrollo de algoritmos adaptativos y protocolos para NCS.

NCSs operan en ambientes estocásticos (azar) en condiciones de incertidumbre. En particular, un nodo sensor, como un tomador de decisiones o el agente, se aplica una acción a su entorno, y luego transita de un estado a otro. El entorno puede abarcar propiedades propias del nodo (por ejemplo, la ubicación de coordenadas y la energía disponible en la batería), así como muchos de los objetos que lo rodean. Por lo tanto, las acciones pueden ser tareas simples o comandos complejos. En un entorno tan incierto, la dinámica del sistema se pueden modelar mediante un marco matemático llamado procesos de decisión de Markov (MDP) para optimizar los objetivos deseados de la red. MDPs implican que el sistema posee una propiedad de Markov. En

particular, el futuro estado del sistema depende sólo del estado actual, pero no los últimos estados. Los acontecimientos recientes en solucionadores MDP han permitido la solución para sistemas de gran escala, y han introducido nuevos potenciales de investigación en NCS.

La estabilidad en el Control y la compensación de retraso

Desde hace muchos años, los investigadores nos han dado las estrategias precisas y óptimas de control que salen de la teoría de control clásica, a partir del control PID, control óptimo, control adaptativo, control robusto, control inteligente y muchas otras formas avanzadas de algoritmos de control. Pero estas estrategias de control deben ser modificados de acuerdo a los requisitos de la aplicación, así como para que funcionen de forma fiable en una red para compensar los retrasos y la imprevisibilidad

La inestabilidad del sistema debido a la demora de la red es un factor muy importante a considerar en la NCS. Se toman diferentes enfoques matemáticos, heurísticas y basados en estadísticas para la compensación de retraso en NCS.

La asignación de ancho de banda y programación

Como hablamos de contar con sistemas multi-actuador de control multi-sensor y en una red, consideración importante se debe dar al ancho de banda disponible en la red. Con la cantidad finita de ancho de banda disponible, queremos utilizarlo de manera óptima y eficiente. Esto plantea la necesidad de que las decisiones de prioridad y cuestiones de programación para controlar una serie de actuadores para una serie de tareas.

Seguridad de Redes

Toda esta discusión sobre el envío de comandos importantes del sensor y de control del actuador en la red nos lleva a un punto importante de la seguridad en la red. Cualquier medio de red inalámbrica es medio especialmente susceptibles a la interceptación fácil; es extremadamente crítica para proteger los datos transmitidos desde el acceso no autorizado y modificaciones en los sistemas inalámbricos. Los usuarios malintencionados pueden interceptar y espiar los datos en tránsito a través de medio compartido y transmitido. Seguridad de la red incluye elementos esenciales en los dispositivos de seguridad de Internet que proporcionan filtrado de tráfico, la integridad, confidencialidad y autenticación. Por lo tanto el intercambio de datos, clasificación de datos y datos / seguridad de la red es de suma preocupación en los sistemas de control en red distribuidos teniendo en cuenta el tiempo y los datos de aplicaciones sensibles.

Bibliografía

- Abu Alsheikh, M., Hoang, D., Niyato, D., Tan, H.-P., & Lin, S. (2015). Markov Decision Processes with Applications in Wireless Sensor Networks: A Survey. *Communications Surveys Tutorials, IEEE, PP(99)*, 1-1.
- Adlakha, S., Lall, S., & Goldsmith, A. (April de 2012). Networked Markov Decision Processes With Delays. *Automatic Control, IEEE Transactions on*, 57(4), 1013-1018.
- Hadjicostis, C. N., & Touri, R. (2002). Feedback control utilizing packet dropping network links. *Decision and Control, 2002, Proceedings of the 41st IEEE Conference on*, 2, págs. 1205-1210.
- Hespanha, J., Naghshtabrizi, P., & Xu, Y. (Jan de 2007). A Survey of Recent Results in Networked Control Systems. *Proceedings of the IEEE*, 95(1), 138-162.
- Inria, E. A. (2004). Constrained Markov Decision Processes.
- Liu, F.-Y. W. (2008). *Networked Control Systems Theory and Applications*. (F.-Y. Wang, Ed.) Springer.
- Monahan, G. E. (1982). State of the art—a survey of partially observable Markov decision processes: theory, models, and algorithms. *Management Science*, 28(1), 1-16.
- Montestruque, L. A., & Antsaklis, P. J. (2003). On the model-based control of networked systems. *Automatica*, 39(10). Obtenido de <http://www.sciencedirect.com/science/article/pii/S0005109803001869>
- Montestruque, L., & Antsaklis, P. (Sept de 2004). Stability of model-based networked control systems with time-varying transmission times. *Automatic Control, IEEE Transactions on*, 49(9), 1562-1572.
- Sigaud, O., & Buffet, O. (2013). *Markov decision processes in artificial intelligence*. John Wiley & Sons.
- Smith, S. C., & Seiler, P. (2003). Estimation with lossy measurements jump estimators for jump systems. *Automatic Control, IEEE Transactions on*, 48(12), 2163-2171.
- Sutton, R. S. (1992). Introduction The challenge of reinforcement learning. En *Reinforcement Learning* (págs. 1-3). Springer.
- Walsh, G., Ye, H., & Bushnell, L. (May de 2002). Stability analysis of networked control systems. *Control Systems Technology, IEEE Transactions on*, 10(3).
- Xu, Y., & Hespanha, J. P. (2005). Estimation under uncontrolled and controlled communications in networked control systems. *Decision and Control, 2005 and 2005 European Control Conference. CDC-ECC'05. 44th IEEE Conference on*, (págs. 842-847).
- Zhang, W., Branicky, M. S., & Phillips, S. M. (2001). Stability of networked control systems. *Control Systems, IEEE*, 21(1), 84-99.