

# Hacia una categorización en el problema de asignación de recursos en logística humanitaria y su resolución utilizando aprendizaje automático

Galo Ruiz-Soto, Miguel González-Mendoza,  
Jaime Mora-Vargas

Instituto Tecnológico y de Estudios Superiores de Monterrey,  
Escuela de Ingeniería y Ciencias,  
México

{A01799399, mgonza, jmora}@tec.mx

**Resumen.** El principal reto de la logística humanitaria es entregar los suministros apropiados en las cantidades adecuadas en el momento preciso y en el lugar correcto, en una situación de emergencia para aliviar el sufrimiento de los supervivientes. Las primeras 72 horas son críticas, no sólo para este fin, sino también para encontrar sobrevivientes que pudieron quedar atrapados después de un desastre natural. Identificamos dos problemas dentro de la logística humanitaria. El primero es el rescate oportuno de sobrevivientes. El segundo problema es la distribución eficiente, efectiva, y justa de los suministros de emergencia. En ambos, identificamos un problema de asignación de recursos y consideramos que el aprendizaje por refuerzo puede ser utilizado para resolverlo. El aprendizaje por refuerzo busca maximizar la recompensa obtenida por acciones en un ambiente en el largo plazo, por lo que es importante para el agente no solo considerar la recompensa inmediata por una acción sino también explorar nuevas acciones esperando obtener el mayor valor. Este es un trabajo exploratorio que busca encontrar una primera categorización del problema de la asignación de recursos y buscamos utilizar técnicas de aprendizaje por refuerzo en las 72 horas críticas posteriores a un desastre natural.

**Palabras clave:** Logística humanitaria, aprendizaje por refuerzo, aprendizaje automático, desastre natural, problema de asignación de recursos.

## Towards a Categorization in the Resource Application Problem in Humanitarian Logistics and its Resolution Using Machine Learning

**Abstract.** The main challenge of Humanitarian Logistics is to deliver the appropriate supplies, in suitable quantities, in the precise moment and the right place in an emergency to relieve human suffering. The first 72 hours are critical not only to achieve this objective but also to find survivors in the aftermath of a natural disaster. We identified two problems in humanitarian logistics. The first one is the rescue of survivors under the rubble after a building collapse. The second problem is the efficient, effective, and equal distribution of emergency

supplies. We have identified a Resource Allocation Problem (RAP) in both cases, and we consider that Reinforcement Learning can be applied to solve the problem. Reinforcement Learning is about maximizing the reward obtained by actions taken in an environment in the long run, therefore it is important for the agent not only to consider the immediate reward but also to explore new actions to maximize the value. This is an exploratory work towards a categorization in the RAP and we intend to use reinforcement learning techniques in the critical 72 hours in the aftermath of a natural disaster.

**Keywords:** Humanitarian logistics, reinforcement learning, machine learning, natural disaster, resource allocation problem.

## 1. Introducción

La logística humanitaria debe de tratarse de manera separada a la logística comercial en varios aspectos, sobre todo tomando en cuenta los objetivos de una logística humanitaria post desastre. La logística humanitaria puede definirse como “un amplio rango de operaciones incluyendo la distribución de suministros médicos para la prevención de enfermedades de rutina, suministros alimenticios para luchar contra el hambre, y suministros críticos en la secuela de un desastre” [3].

Sin embargo, otros factores podrían considerarse en la definición tales como la velocidad de la distribución de los recursos, la asignación de políticas implementadas por las autoridades y la necesidad de encontrar sobrevivientes después de un desastre. “Cuando un desastre mayor golpea, una respuesta oportuna para salvar vidas y mitigar los sufrimientos de la población afectada se vuelve crítica.

De hecho, las primeras 72 horas de un esfuerzo de alivio de un desastre son críticas ya que las oportunidades de supervivencia posteriores a esa ventana de tiempo sin agua ni comida decrecen drásticamente. El reto es entregar los suministros apropiados de emergencia en cantidades suficientes exactamente cuándo y dónde son necesarios” [11].

El Problema de la Asignación de Recursos (RAP por sus siglas en inglés) en la logística humanitaria plantea que “el objetivo es entregar los recursos escasos de forma que se maximice la efectividad de la ayuda” [6].

La investigación se dirige hacia resolver el RAP utilizando técnicas de aprendizaje automático y en particular se explorará el aprendizaje por refuerzo y se contrastarán los resultados con otros modelos. La Fig. 1 presenta una representación gráfica del RAP. Es posible que se reciban donaciones (medicinas, agua, alimentos no perecederos, cobijas) de particulares, empresas u organizaciones. Estas donaciones se llevan a un centro de recursos para su clasificación y de ahí se distribuyen a refugios temporales.

De acuerdo con Sutton y Barto [9], el aprendizaje por refuerzo se refiere a aprender qué hacer para obtener la mayor recompensa. Sin embargo, no existe una guía para indicarle a un agente qué hacer, sino que debe de descubrirlo por sí mismo pensando que sus acciones pueden afectar no sólo la recompensa más inmediata sino también las recompensas siguientes. “Estas dos características, ensayo y error y recompensa retrasada, son las dos características distintivas más importantes del aprendizaje por refuerzo” [9].

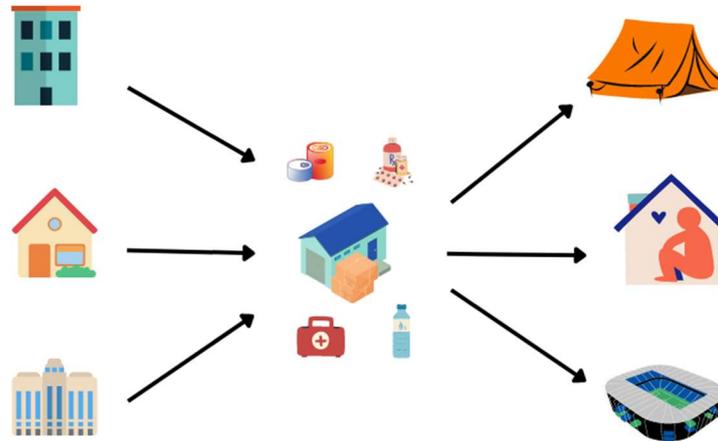


Fig. 1. Representación general del problema de la asignación de recursos en una situación de desastre.

Sin embargo, de acuerdo con Sutton y Barto, hay un juego entre explotar y explorar ya que “el agente tiene que explotar lo que ha experimentado anteriormente para obtener recompensa, pero también debe de explorar para lograr mejores elecciones en el futuro.

El agente debe de intentar una variedad de acciones y progresivamente favorecer aquellas que aparentemente son las mejores... Todos los agentes de aprendizaje por refuerzo tienen objetivos específicos, pueden sentir aspectos de su entorno, y pueden escoger acciones para influir en sus ambientes. Además, se suele asumir desde el inicio que el agente tiene que operar a pesar de una incertidumbre significativa sobre el ambiente al que se enfrenta” [9].

En el aprendizaje por refuerzo se cuentan cuatro elementos, una política, una señal de recompensa, una función de valor además de un modelo del ambiente. La política habla del comportamiento de un agente en un momento dado. La señal de recompensa define los buenos y malos eventos para el agente por lo que la política puede modificarse. En cada etapa el agente recibe un número que será conocido como la recompensa, haciendo que el objetivo del agente sea maximizar la recompensa total en el largo plazo.

Los valores indican el atractivo de los estados del ambiente después de tomar en cuenta aquellos estados que pueden seguir y las recompensas que se pueden encontrar ahí. Los valores son importantes para lograr mayor recompensa. En la evaluación de las decisiones se buscan acciones que provean el mayor valor y no la mayor recompensa dado que estas acciones obtendrán el mayor monto de recompensa sobre el tiempo [9].

## 2. Trabajos relacionados

Fan et al. [1] utilizan Deep Q Network (DQN) para la distribución de suministros a gran escala. Consideran centros locales de recursos que se encargan de distribuir

suministros a las distintas áreas afectadas (AA), sin embargo, opinan que las necesidades de recursos pueden ser distintas en cada una de las AA.

Su objetivo es “obtener la política óptima de distribución de suministros de emergencia que minimice el costo económico mientras se minimiza el sufrimiento de los sobrevivientes” [1]. Utilizan tres criterios, la eficiencia reflejada en el costo de transporte de los suministros, la efectividad manifestada en el costo de la privación de los suministros de emergencia, y la equidad reflejada por el costo de la distribución injusta de los suministros.

Yu et al. [10] también consideran las métricas de eficiencia (el costo de entrega basado en accesibilidad), efectividad (fundamentada en el costo inicial de los suministros recibidos además de la demanda), y equidad (reflejada en el costo de que todos los sobrevivientes obtengan las mismas condiciones), aplicando un algoritmo de Q-learning.

Sin embargo, consideran refugios temporales como AA y “sirven como el punto donde se reciben los recursos desde fuera del AA y se distribuyen a los supervivientes al interior del AA” [10]. Toman en cuenta un Centro Local de Respuesta (CLR) que recibe recursos desde un centro mayor de distribución usualmente administrado por el gobierno local. Utilizan accesibilidad en lugar de distancia para medir el costo de entrega [10]. De acuerdo con su investigación, su algoritmo de Q-learning supera al algoritmo de programación dinámica.

A pesar de que Hachiya et al. [2] utilizan técnicas de aprendizaje por refuerzo para la distribución de suministros de emergencia, solo consideran vehículos aéreos no tripulados (UAV, por sus siglas en inglés), los cuales pueden ser muy útiles en el transporte en caso de que vías terrestres estén bloqueadas después de un desastre, sin embargo, su capacidad de carga es relativamente pequeña por lo que para hacer una distribución justa de los suministros se tendrían que realizar varios viajes.

Mora-Ochomogo et al. [5] aplican el proceso de decisión de Markov para representar las operaciones en CLR que reciben donaciones. Los CLR deciden qué donaciones recibir por lo que el costo de equidad no se aplica de la misma manera que lo hacen otros autores. La tabla 1 presenta un resumen de trabajos relacionados a la resolución del RAP con distintos modelos o marcos de trabajo.

### **3. Hacia una categorización en el problema de asignación de recursos en logística humanitaria**

Se pueden identificar dos problemas distintos en logística humanitaria. Por un lado, es esencial poder transportar suministros de emergencia a las áreas afectadas. Es una tarea de la “logística humanitaria explorar una estrategia eficiente, efectiva y justa de distribución de suministros para reducir el sufrimiento humano causado por la escasez de materiales en las áreas influenciadas” [1]. Este es usualmente el problema que se considera en el RAP.

Sin embargo, hay un segundo flujo que es el rescate de sobrevivientes que puedan estar atrapados bajo los escombros después del colapso de un edificio o estructura. Dado que los recursos son escasos, los rescatistas deben de decidir en qué sitios poner los esfuerzos de rescate para encontrar sobrevivientes.

**Tabla 1.** Trabajos relacionados.

Referencia	Problema	Modelo/Marco	Observaciones
Yu et al. (2021)	Asignación de recursos	Q-learning	Requerimientos iguales en cada AA
Fan et al (2022)	Asignación de recursos	DQN	Requerimientos diferentes en cada AA
Hachiya et al. (2022)	Asignación de recursos	Q-learning	Vehículos aéreos no tripulados
Mora-Ochomogo et al. (2022)	Asignación de recursos	Proceso de decisión de Markov	No hay costo de equidad

Las primeras 72 horas son cruciales para la oportunidad de rescatar gente debajo de los escombros y asignar los recursos críticos como agua y comida para los sobrevivientes. También deben de tomarse en cuenta otro tipo de recursos como pueden ser palas mecánicas para remover escombros de manera más rápida. Consideramos que este segundo problema ha sido menos estudiado en la literatura y debería de ser considerado en el RAP.

Mencionamos anteriormente que el RAP en la logística humanitaria indica que se busca distribuir los recursos maximizando la efectividad de la ayuda. En el caso de la presente investigación, se consideran distintas AA después de un desastre natural por lo que el objetivo es maximizar la efectividad de la ayuda, ya sean suministros o rescatistas, en cada una de las AA.

Sin embargo, consideramos, al igual que Fan et al. [1] que las necesidades de recursos en las AA no tienen por qué ser las mismas en cada una de ellas. El número de edificaciones afectadas después de un desastre natural puede hacer variar las necesidades de recursos.

La técnica que se pretende utilizar para atacar el RAP es aprendizaje por refuerzo. Se realizó una primera categorización del RAP, encontrando dos problemas distintos en la logística humanitaria, y ambos pueden ser abordados con aprendizaje por refuerzo.

El primer problema en nuestra categorización, identificado en la logística humanitaria, es el rescate de sobrevivientes que puedan estar atrapados bajo los escombros después del colapso de un edificio o estructura.

En este caso, consideramos que existe un RAP ya que se deben de considerar recursos que deben de ser entregados de manera eficiente y justa buscando maximizar la efectividad de la ayuda y facilitar el rescate de sobrevivientes. Estos recursos pueden incluir un número adecuado de rescatistas entrenados, perros preparados para rescate, grupos de topes, palas mecánicas para remover escombros, herramienta especializada, entre otros.

Consideramos que este primer problema puede ser abordado con aprendizaje por refuerzo. En los últimos años, la sociedad civil ha contribuido de manera importante en el rescate de sobrevivientes en edificios colapsados.

La sociedad civil se auto organiza para establecer centros de rescate no gestionados por el gobierno. Por lo tanto, existen dos vertientes que buscan lograr el mismo fin, pero la gestión es distinta. Los agentes estarán buscando la asignación de recursos dedicados a rescatar sobrevivientes.

**Tabla 2.** Problemas identificados en la logística humanitaria.

Problema	Subproblema	Agentes involucrados
Estrategia eficiente, efectiva y justa de distribución de suministros	1) Distribución de recursos a una AA después del colapso de un edificio	Donantes de recursos
	2) Distribución de recursos a un albergue temporal	Distribuidores y donantes de recursos
Rescate de sobrevivientes		Rescatistas, tomadores de decisiones.

La recompensa que obtendrán de manera inmediata será el rescate de personas bajo los escombros, sin embargo, hay agentes que tendrán que buscar maximizar el valor en el tiempo dependiendo de los recursos escasos dedicados a tareas de rescate. Un ejemplo de lo anterior es la asignación de una pala mecánica para remover escombros dado que es un recurso muy escaso y debería de ser asignada a una AA donde se pueda maximizar el número de sobrevivientes rescatados. El segundo problema que se categorizó habla de encontrar una estrategia eficiente y efectiva de distribución de suministros de emergencia.

Sin embargo, consideraremos que la distribución no deberá de ser justa en el sentido de que todos los CRL recibirán la misma cantidad de insumos, sino que las necesidades serán diferentes en cada uno de ellos dado que la situación de emergencia seguramente será distinta en cada AA. Por ejemplo, dependerá de si hubo un colapso de edificio en el caso de un temblor, o si hay un refugio que está recibiendo gente que no puede volver a casa. En el caso de un colapso de un edificio, usualmente la sociedad civil se organiza para recibir suministros de emergencia tanto para posibles afectados por el derrumbe, pero también se recibe comida preparada o bebidas para los rescatistas que estarán en busca de sobrevivientes debajo de los escombros.

Usualmente la literatura considera un único centro de recursos centralizado (CRC) el cual se encargaría de distribuir suministros de emergencia a las AA. Sin embargo, debemos de analizar más a fondo si atacar el RAP con este mismo supuesto o considerar un CR local (CRL) en cada una de las AA debido a la propia organización de la sociedad civil ya que los suministros pueden ser enviados de manera directa al CRL sin pasar por un CRC, usualmente manejado por el gobierno.

En este caso que estamos considerando, los agentes no recibirán instrucciones de qué suministros ni las cantidades que tendrán que llevar al CRL, ni tampoco de cómo transportar los insumos. Usualmente estos agentes llevan los insumos a los CRL cercanos a su domicilio. Recibirán recompensas inmediatas en caso de que el CRL acepte los suministros que estén llevando

Sin embargo, llegará un momento donde el CRL no pueda recibir ya ciertos insumos por lo que los agentes dejarán de recibir recompensa. Mientras tanto, otros agentes

estarán dispuestos a explorar y asistir a un CRL distinto o incluso al CRC obteniendo una recompensa tardía, pero haciendo mayor el valor en el tiempo.

En el caso de personas afectadas por un desastre natural y que no quedaron atrapadas bajo los escombros de un edificio serían trasladadas a un refugio temporal para recibir asistencia médica y alimentos. En este caso, usualmente se reciben donaciones de particulares, organizaciones o empresas y éstas se llevan a un CRC.

En el centro de recursos se clasifican y posteriormente se distribuyen a los refugios, ya sean centralizados o en las AA. Este caso es similar al anterior, sin embargo, se agregan agentes que requieren distribuir los recursos a los refugios temporales dentro del período crítico. Es posible algunos caminos para llegar del CRC a los refugios se encuentren bloqueados por escombros, o incluso pueda haber réplicas del desastre natural que compliquen la tarea de distribución.

Los agentes no recibirán instrucciones de qué caminos tomar o cómo entregar los suministros de emergencia, pero recibirán recompensas inmediatas si entregan dentro del período crítico de 72 horas, incluso si transitan por áreas peligrosas o con obstáculos en el camino. Algunos de estos agentes estarán dispuestos a explorar rutas sostenibles en el tiempo para poder distribuir los recursos de manera más eficiente a los albergues temporales, retardando así la recompensa, pero el valor pudiera ser maximizado en el tiempo.

#### **4. Conclusiones y trabajo a futuro**

En el presente trabajo revisamos el problema ya clásico de la logística humanitaria que es entregar los suministros apropiados en las cantidades adecuadas en el momento preciso y en el lugar correcto en una situación de emergencia para aliviar el sufrimiento de los supervivientes en las primeras 72 horas críticas. Para lograr ese fin, se debe de abordar el problema de asignación de recursos.

Sin embargo, pretendemos establecer una categorización del problema ya que hemos distinguido dos vertientes importantes y consideramos que una de ellas debe de ser considerada como prioritaria. Por un lado, está el rescate de sobrevivientes y que no ha sido abordado de manera extensa en la literatura. Por otro lado, está la distribución de suministros de emergencia a las áreas afectadas o a albergues temporales y que ha sido abordado de manera más extensa en la literatura.

Pretendemos aplicar técnicas de aprendizaje por refuerzo donde los agentes estarán en busca de recompensas inmediatas como lo es encontrar sobrevivientes o entregar suministros suficientes en las primeras horas críticas posteriores a un desastre natural. Dado que los recursos son escasos, es crucial la búsqueda de recompensas tardías para maximizar el valor en el tiempo. Este es un trabajo exploratorio y hemos presentado el planteamiento de la investigación que pretendemos realizar en el corto plazo.

En el presente trabajo presentamos una categorización del problema de la asignación de recursos y pretendemos atacar de manera inicial el rescate de sobreviviente. En una segunda etapa abordaremos el problema de entrega de suministros de emergencia. Como trabajo a futuro, consideramos que podemos extender la categorización con nuevas vertientes del problema conforme avance nuestra investigación.

## Referencias

1. Fan, J., Chang, X., Mišić, J., Mišić, V. B., Kang, H.: DHL: Deep reinforcement learning-based approach for emergency supply distribution in humanitarian logistics. *Peer-to-Peer Networking and Applications*, vol. 15, no. 5, pp. 2376–2389 (2022) doi: 10.1007/s12083-022-01353-0
2. Hachiya, D., Mas, E., Koshimura, S.: A reinforcement learning model of multiple UAVs for transporting emergency relief supplies. *Applied Sciences*, vol. 12, no. 20, pp. 10427 (2022) doi: 10.3390/app122010427
3. Holguín-Veras, J., Jaller, M., Wassenhove, L. N. V., Pérez, N., Wachtendorf, T.: On the unique features of post-disaster humanitarian logistics. *Journal of Operations Management*, vol. 30, no. 7-8, pp. 494–506 (2012) doi: 10.1016/j.jom.2012.08.003
4. Holguín-Veras, J., Pérez, N., Jaller, M., Wassenhove, L. N. V., Aros-Vera, F.: On the appropriate objective function for post-disaster humanitarian logistics models. *Journal of Operations Management*, vol. 31, no. 5, pp. 262–280 (2013) doi: 10.1016/j.jom.2013.06.002
5. Mora-Ochomogo, I., Serrato, M., Mora-Vargas, J., Akhavan-Tabatabaei, R.: Application of a Markov decision process in collection center operations. *Humanitarian Logistics from the Disaster Risk Reduction Perspective*, pp. 407–428 (2022) doi: 10.1007/978-3-030-90877-5\_14
6. Pérez-Rodríguez, N., Holguín-Veras, J.: Inventory-allocation distribution models for postdisaster humanitarian logistics with explicit consideration of deprivation costs. *Transportation Science*, vol. 50, no. 4, pp. 1261–1285 (2016) doi: 10.1287/trsc.2014.0565
7. Sheu, J.: An emergency logistics distribution approach for quick response to urgent relief demand in disasters. *Transportation Research Part E: Logistics and Transportation Review*, vol. 43, no. 6, pp. 687–709 (2007) doi: 10.1016/j.tre.2006.04.004
8. Sphere: Humanitarian charter and minimum standards in disaster response. Technical report, The Sphere Project, Geneva, Switzerland (2018) [www.spherehandbook.org/](http://www.spherehandbook.org/)
9. Sutton, R. S., Barto, A. G.: Reinforcement learning: An introduction (Adaptive computation and machine learning). The MIT Press, Cambridge, 2nd Edition (1998)
10. Yu, L., Zhang, C., Jiang, J., Yang, H., Shang, H.: Reinforcement learning approach for resource allocation in humanitarian logistics. *Expert Systems with Applications*, vol. 173, pp. 114663 (2021) doi: 10.1016/j.eswa.2021.114663
11. Zeimpekis, V., Ichoua, S., Minis, I.: Humanitarian and relief logistics. *Operations Research/ Computer Science Interfaces Series* (2013) doi: 10.1007/978-1-4614-7007-6