

Método de Reducción de Incertidumbre basado en Algoritmos Evolutivos Paralelos orientado a Procesos de Predicción

Bianchini Germán¹, Caymes-Scutari Paola^{1,2}, Méndez-Garabetti Miguel^{1,2,3},
Tardivo María Laura^{1,2,4}

¹Laboratorio de Investigación en Cómputo Paralelo/Distribuido
Departamento de Ingeniería en Sistemas de Información
Facultad Regional Mendoza/Universidad Tecnológica Nacional
Rodríguez 273 (M5502AJE) Mendoza, +54 261 5244579

²Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET)

³Instituto Tecnológico Universitario, Universidad Nacional de Cuyo, Centro
Universitario, Mendoza, Argentina

⁴Departamento de Computación, Universidad Nacional de Río Cuarto, Río Cuarto,
Córdoba, Argentina

gbianchini@frm.utn.edu.ar, pcaymesscutari@frm.utn.edu.ar, miguelmendezgarabetti@gmail.com,
lauratardivo@dc.exa.unrc.edu.ar

Resumen

El modelado matemático, la simulación, la visualización, la programación y la computación de alto rendimiento, propician el estudio de problemas que resultan muy complejos, demandantes de tiempo y espacio (en términos de memoria y almacenamiento), e incluso muy peligrosos por los riesgos humanos y/o ambientales que conllevan. Sin embargo, aún cuando la formulación del modelo y el algoritmo de implementación estén libres de errores, es muy difícil que el simulador ofrezca una salida fiel a la situación real que se desea representar dada la incertidumbre introducida tanto por las expresiones que pueden no responder a un modelo estrictamente matemático como por las limitaciones de la representación numérica inherente a la arquitectura del hardware subyacente. De esta manera, la reducción de incertidumbre constituye un

proceso muy importante para lograr predicciones más precisas. Un problema concreto dentro de las limitaciones de los modelos suele estar relacionado con la gran cantidad de parámetros de entrada que deben manejar. Dichos parámetros suelen presentar algún tipo de incertidumbre debido a la imposibilidad de medirlos en tiempo real, y por lo tanto es común que deba recurrirse a la utilización de estimaciones producto de medidas indirectas. El proyecto aquí comentado propone el diseño y desarrollo de un método, basado en Algoritmos Evolutivos Paralelos, HPC y Estadística, para reducir la incertidumbre de los parámetros de entrada en procesos de predicción de propagación.

Palabras clave: HPC, Reducción de Incertidumbre, Cómputo Paralelo, Algoritmos Evolutivos, Modelo de Islas, Predicción.

Contexto

El proyecto se encuentra en ejecución desde enero de 2014 en el LICPaD (Laboratorio de Investigación en Cómputo Paralelo/Distribuido) en la UTN-FRM, y forma parte de la continuación de los proyectos PID UTN1194 (de la UTN) y PIP 11220090100709 (del CONICET), en los cuales se desarrolló el método ESS (*Evolutionary-Statistical System*) [1]. Dicho método involucra algoritmos evolutivos [2] y se basa en un esquema de única población con evaluación paralela de la aptitud [3]. En contraste, el método propuesto en el proyecto que aquí se trata está enfocado en la utilización del mismo esquema evolutivo, pero potenciando la utilización de los recursos paralelos, a partir de la operación sobre múltiples poblaciones y utilización de un esquema de migración (conocido como Modelo de Islas) [4].

El presente proyecto se encuentra financiado por la UTN, a través del proyecto EIUTIME0002169TC, y por el MINCyT mediante el proyecto PRH PICT 2008-00242.

Introducción

Dentro de la presente línea de I+D, hemos desarrollado como antecedentes dos métodos que se relacionan con la temática del proyecto tratado en el presente artículo. El primero, basado en los conceptos de experimentación factorial, análisis estadístico y cómputo paralelo, denominado S^2F^2M (*Statistical System for Forest Fire Management*) [5], se centra en hallar un patrón del comportamiento del modelo en el cual se aplica, a través de la generación de los posibles escenarios (combinaciones de valores de parámetros de entrada del

modelo) y la evaluación del conjunto de resultados para ofrecer una tendencia en el comportamiento del modelo, ajustándolo con la observación actual y real del mismo. El segundo, realizado en el marco del PID UTN1194 y denominado ESS (*Evolutionary-Statistical System*) [1], constituye un método conducido por datos de múltiples soluciones solapadas, donde cada solución se representa a través de un mapa basado en celdas. Por un lado, ante la incertidumbre existente en los parámetros de entrada del modelo, ESS requiere que para cada parámetro se indique un dominio, de forma tal que pueda variarse el valor del parámetro dentro de dicho dominio. De esta manera, las posibles combinaciones en el conjunto de valores de entrada del modelo resultan numerosas, de las cuales algunas podrán describir la realidad de forma más fiel que otras. Cada interpretación particular de los parámetros del modelo se denomina 'escenario', y a su vez puede considerarse un 'individuo' en el paradigma evolutivo que es el utilizado en el método. Por otro lado, ESS es un método iterativo en el sentido de que la predicción del fenómeno se realiza cada cierto intervalo de tiempo, considerando el estado anterior y actual de la propagación. La observación del estado real se utiliza para calibrar la predicción y lograr predicciones más ajustadas para los pasos de simulación siguientes. En cada iteración, ESS presenta dos fases generales: la fase de optimización (OS u *Optimization Stage*) que permite orientar la búsqueda hacia los escenarios que resulten más fieles a la realidad (implementada mediante un algoritmo evolutivo paralelo [3]), y la fase de Calibración (CS o *Calibration Stage*) encargada de establecer la tendencia y proveer una predicción (fase

implementada a través de un método estadístico).

En cada caso, los resultados de la aplicación de ambos métodos sobre casos de incendios forestales reales, ha ofrecido muy buenos valores, superando los resultados alcanzados con la aplicación de otros métodos equivalentes (para resultados comparativos puede consultarse [1,6,7]).

Dentro de los objetivos actuales se tiene el desarrollo de diversas variantes de ESS con aplicación de paralelismo a distintos niveles, ya que el desarrollo de ESS aplica el esquema paralelo a nivel de función de evaluación (*fitness function*). En el proyecto aquí comentado se ha planteado un esquema de islas [4] en el cual la población de individuos se vea distribuido en diversas subpoblaciones que se administren y puedan evolucionar de forma paralela. Sin embargo, esta nueva implementación (que utiliza el paradigma *Master-Worker* [8,9]) utiliza dos niveles: **Master-WorkerL1** y **Master-WorkerL2**, tal como puede observarse en la Figura 1.

Por un lado, en el nivel **Master-WorkerL1**, el rol de *MasterL1* es asumido por un proceso denominado Monitor que inicializa cada una de las islas (o *WorkersL1*) al comienzo de la ejecución, luego recibe los resultados de cada una de las islas y por último lleva a cabo la predicción del comportamiento del incendio para cada paso de simulación (es decir, los sucesivos instantes de tiempo en los que se realiza la predicción). Por otro lado, el nivel **Master-WorkerL2**, está compuesto por diferentes instancias de este esquema, donde cada isla contiene un proceso *MasterL2* y su respectivo grupo de *WorkersL2*. El *MasterL2* de cada isla distribuye los individuos a los *WorkersL2*, donde un individuo en ESS-IM (como ya se comentó para el caso de

ESS) representa un conjunto particular de parámetros de entrada que alimentan al modelo.

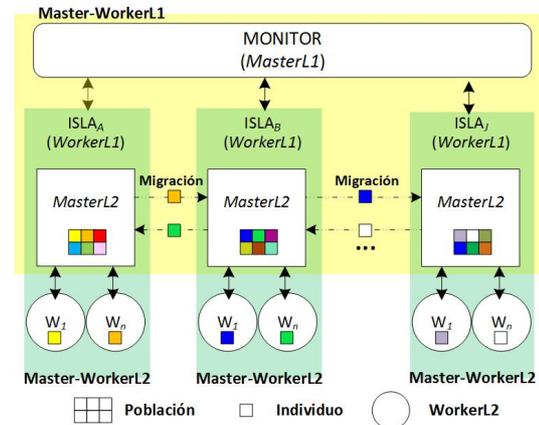


Figura 1. Modelo de comunicación ESS-IM

Los *WorkersL2*, por su parte, llevan a cabo la simulación y la evaluación de la aptitud de cada uno de los individuos en paralelo, retornando los resultados a su respectivo *MasterL2*. Cada *MasterL2*, además, es responsable de llevar a cabo la evolución de la población y de efectuar las migraciones de los individuos entre las distintas islas, en función de la topología de comunicación y los parámetros de migración configurados. Es importante remarcar que cada isla maneja una población independiente y que la evolución en las diferentes islas se efectúa en paralelo.

Líneas de Investigación, Desarrollo e Innovación

La línea de I+D en la que se enmarca el presente proyecto y los precedentes mencionados en el apartado Contexto, abarca el análisis, diseño y desarrollo de métodos para tratar la incertidumbre en los valores de los parámetros de entrada de modelos ambientales como lo es el caso concreto de los incendios forestales.

Dichos modelos centran sus bases en conceptos tales como el análisis estadístico, el cómputo paralelo y la aplicación de metaheurísticas. Los métodos desarrollados se hallan clasificados dentro de los Métodos Conducidos por datos con Solución Solapada Múltiple (DDMMOS, por su sigla en inglés), dado que la solución que aportan no se basa en el resultado de un único caso sino en la conjunción de una multiplicidad de casos estudiados. Su objetivo es hallar un patrón de comportamiento, independientemente de los valores de los parámetros del modelo en cuestión. En estos métodos, cada parámetro se representa como un conjunto de valores dentro de un rango válido. Luego se generan múltiples escenarios a través de diversas combinaciones de los valores de los parámetros de entrada (valores generados de forma aleatoria y calibrados a través de una metaheurística), y entonces se evalúa el modelo para cada caso. Los resultados son agregados estadísticamente para determinar la probabilidad general. En el caso particular de los incendios forestales, esta agregación se utiliza para predecir el área quemada en sucesivos pasos de simulación.

Resultados y Objetivos

Como resultado de este desarrollo se ha logrado, en primer lugar, obtener mejoras en la calidad de la predicción alcanzada por el modelo (a raíz de la mayor diversidad de casos lograda mediante la multiplicidad de poblaciones), y en segundo término, mejorar los tiempos de ejecución de la metodología. Para verificar tales mejoras, los resultados de ESS-IM se han comparado con los producidos por las versiones anteriores (ESS y S²F²M) a

través de una serie de experimentos que involucraron la utilización de datos de incendios reales, y que fueron publicados en [1, 6, 7, 10].

Actualmente se está trabajando en otras modificaciones, que redunden en mejoras de diversos aspectos del método (como por ejemplo la utilización de Evolución Diferencial como metaheurística o la variación dinámica de la representación del terreno) y en su aplicación sobre distintas estructuras de hardware (por ejemplo la adecuación para su operación sobre GPUs).

Formación de Recursos Humanos

El desarrollo del presente proyecto favorecerá la consolidación del grupo de investigación ya que, desde su creación a la fecha, ha ido incorporando investigadores y su producción científica se ha visto incrementada directamente. Aparte de la innovación científica, se espera que el presente proyecto provea de medios económicos para el fortalecimiento de las líneas de investigación relacionadas y brinde soporte a las actividades que se desprenden del proyecto aquí planteado. Tales líneas se extenderán mediante la incorporación de nuevos becarios y estudiantes avanzados, principalmente de la carrera de Ingeniería en Sistemas de Información.

Actualmente, a nivel de grado, el laboratorio cuenta con dos alumnos activos de la carrera Ingeniería en Sistemas de Información, y dos graduados. Por su parte, a nivel de postgrado, dentro del grupo de investigación y del presente proyecto, contamos con la participación del Ing. Miguel Méndez Garabetti, quien cursa su formación doctoral en la Universidad Nacional de San Luis y desarrolla su plan

de tesis enmarcado en una temática afín con el presente proyecto, bajo la dirección del Dr. Germán Bianchini en el marco del LICPaD. Por su parte, la Lic. María Laura Tardivo también está realizando sus estudios de doctorado en la UNSL bajo la dirección de la Dra. Paola Guadalupe Caymes Scutari, quien es Co-Directora del presente proyecto.

Referencias

- [1] G. Bianchini, P. Caymes Scutari, M. Méndez Garabetti, Evolutionary-Statistical System: a Parallel Method for Improving Forest Fire Spread Prediction, *Journal of Computational Science (JOCS, Elsevier)*. Vol. 6, pp. 58-66, 2015.
- [2] D. E. Goldberg, Genetic and evolutionary algorithms, *Come of age*, *Communications of the ACM*, 37 (3) pp. 113-119, 1994.
- [3] E. Cantú Paz, A survey of Parallel Genetic Algorithms. *Calculateurs Parallèles, Réseaux et Systems Repartis*, 10(2), pp. 141-171, 1998.
- [4] Talbi, El-Ghazali, "Metaheuristics – From Design to Implementation", ISBN 978-0-470-27858-1. Wiley. 2009.
- [5] G. Bianchini, A. Cortés, T. Margalef, E. Luque, S²F²M – Statistical System for Forest Fire Management, *LNCS 3514*, pp. 427-434, 2005.
- [6] G. Bianchini, M. Denham, A. Cortés, T. Margalef, E. Luque, Wildland Fire Growth Prediction Method Based on Multiple Overlapping Solution, *Journal of Computational Science* Vol. 1 Issue 4, pp. 229-237, 2010.
- [7] G. Bianchini, M. Méndez Garabetti, P. Caymes Scutari, Evolutionary-Statistical System for Uncertainty Reduction Problems in Wildfires, XVIII CACIC, ISBN: 978-987-1648-34-4, pp. 230-238, 2012.
- [8] A. Grama, A. Gupta, G. Karypis, V. Kumar, "Introduction to Parallel Computing", Second Edition. Pearson, 2003.
- [9] T. Mattson, B. Sanders, B. Massingill, "Patterns for Parallel Programming", Addison-Wesley, 2005.
- [10] M. Méndez Garabetti, G. Bianchini, M. Tardivo, P. Caymes Scutari, Predicción del Comportamiento de Incendios Forestales mediante aplicación de Modelo de Islas a 'ESS', *Noveno Encuentro del "International Center for Earth Sciences" (E-ICES9)*, Actas trabajos completos E-ICES 9, 1a Ed. ISBN 978-987-1323-36-4. Edit. CNEA (Bs.As). pp. 232-246, 2014.