

Modelos para anticipar la factibilidad de que un proyecto de desarrollo de software sea trazable

Juan Giró¹, Juan Vázquez¹, Brenda Meloni¹, Leticia Constable¹ y Andrea Jorner¹

¹ Departamento de Ingeniería en Sistemas de Información
Facultad Regional Córdoba, Universidad Tecnológica Nacional
Maestro López esq. Cruz Roja Argentina, Córdoba, juanfgiro@gmail.com

Resumen. Pese al tiempo transcurrido desde el nacimiento de la ingeniería de software y al enorme esfuerzo realizado, hay pocas evidencias de que los progresos en el campo de la trazabilidad trasciendan los ámbitos de investigación y sean efectivamente aplicados en la industria del software. Esto nos estimuló a buscar las causas de este problema y a proponer modelos que permiten anticipar la factibilidad de que un proyecto sea trazable. Estos modelos se apoyan en ciertos factores considerados determinantes y en métricas que permiten cuantificarlos. Se presentan dos modelos, uno básico apoyado en “diagramas de radar” y otro avanzado que es implementado con una red neuronal de perceptrones multicapa. El ajuste de estos modelos se apoya en información de la industria y para ello se realizaron tareas de relevamiento, que han permitido disponer de los primeros resultados.

Keywords: ingeniería de software, desarrollo de sistemas, trazabilidad.

1 Introducción

Ya han transcurrido más de cuarenta años desde que nació (año 1968) la Ingeniería de Software en cuyo contexto se establecieron numerosas áreas de investigación y desarrollo que, a través de métodos, técnicas y herramientas, han venido haciendo sustanciales aportes a la industria del software y donde una de las principales áreas de estudio es la trazabilidad. Al hablarse de trazabilidad en los proyectos de desarrollo de software se está haciendo referencia a la gestión de un hilo conductor que vincula a las numerosas etapas de su ciclo de vida, brindando la necesaria garantía de coherencia, completitud y corrección del software producido y posibilitando su eficaz mantenimiento correctivo y preventivo.

Sin embargo, pese al tiempo transcurrido y al enorme esfuerzo realizado, hay pocas evidencias de que los progresos en el campo de la trazabilidad trasciendan los ámbitos de investigación y sean efectivamente aplicados en la industria del software, en concordancia con las cualidades que pregonan sus autores y los indudables beneficios que aportarían tanto a los proyectos como a sus productos. La persistencia de dificultades en la industria del software ya fue expuesta en términos generales por Gibbs en su célebre artículo “Software's Chronic Crisis” [1], y una referencia específica a las causas de los

problemas de la trazabilidad fue presentada por Gotel y Finkelstein [2]. Mucho más recientemente, Rilling et al. [3] y Cleland-Huang [4] analizaron el panorama actual y sus perspectivas futuras.

El presente trabajo está inspirado en la necesidad de identificar, en el estado actual del arte, las causas que obstaculizan la aplicación efectiva de la trazabilidad en la industria del software con un doble objetivo: *i*) poder hacer recomendaciones a la industria para superar estas dificultades y *ii*) reconocer las deficiencias o limitaciones de los recursos disponibles y que deberían ser considerados por los investigadores en futuras propuestas. El artículo está organizado de la siguiente manera: en la sección 2 se enuncia la idea presentada, en la sección 3 se describen los modelos destinados a respaldar la propuesta y en la sección 4 se exhiben y discuten los resultados obtenidos hasta el momento. Finalmente en la sección 5 se incluyen las conclusiones de este trabajo y previsiones para su continuación.

2 Descripción de la Propuesta

Como ya fue anticipado, se advierte que los recursos surgidos de la Ingeniería de Software en el campo de la trazabilidad presentan dificultades a la hora de ser aprovechados en la industria y estas dificultades tienen diversas manifestaciones. Entre otros problemas, puede citarse que la trazabilidad *i*) no acompaña a los proyectos en todo el ciclo por ser abandonada en una etapa temprana y que *ii*) es implementada de manera muy superficial o que es aplicada sólo en algunas fases contiguas en el ciclo de vida de los productos software. Aquí se comprueba una incongruencia entre lo proclamado (de manera casi unánime se reconocen los beneficios y la conveniencia de acompañar un proyecto mediante un proceso de trazabilidad) y lo verdaderamente realizado en la práctica habitual.

Habiéndose expuesto el problema, surge la presunción de que no es fortuito que un cierto proyecto pueda ser exitosamente trazado y otro no. Por el contrario, seguramente hay una combinación de condiciones, absolutamente deterministas, que conducen a uno u otro resultado. De ser así, debería ser posible identificar los factores que determinan si un cierto proyecto de desarrollo de software podrá ser, o no, efectivamente respaldado por un sistema de trazabilidad.

Al plantearse como hipótesis la existencia de tales factores y la posibilidad de identificarlos, es necesario hacer las siguientes consideraciones: *i*) estos factores estarán necesariamente vinculados al “estado del arte” de la ingeniería de software en cada momento, es decir que serán cambiantes en el tiempo, *ii*) en base a estos factores debería ser posible anticipar la conveniencia de aplicar trazabilidad a cierto proyecto concreto, o la conveniencia de no hacerlo en ciertas condiciones y en un momento dado.

Quedan así establecidos los objetivos del trabajo presentado, que son: *i*) identificar los factores determinantes en el uso efectivo de la trazabilidad en los proyectos de desarrollo de software, *ii*) proponer modelos que reúnan estos factores y permitan anticipar los resultados que se esperan de cada caso y *iii*) ajustar y evaluar estos modelos a partir de información relevada en la industria del medio.

2.1 Entidades, Dimensiones y Factores

La actividad de la industria del software está relacionada con tres entidades principales que están estrechamente relacionadas entre sí, que son: a) Producto, b) Proyecto, c) Organización. Por lo tanto, parece natural anticipar que los factores buscados deben estar asociados a dimensiones de estas entidades, estableciéndose una escala jerárquica que reconoce entidades, dimensiones y factores. En algunos casos los factores son las propias dimensiones y en otros son una combinación de dimensiones que tiene por finalidad reducir la cantidad de variables, “parámetros característicos” o “grados de libertad” del problema. Conceptualmente estos factores deben ser independientes unos de otros, es decir ortogonales entre sí, y deben tener asignados una unidad de medida y escala de evaluación.

Los factores a ser adoptados deben ser de gran impacto en el problema tratado y en la Tabla 1 se presentan los elegidos en esta propuesta.

Tabla 1. Definición de entidades, dimensiones y factores característicos.

Entidad	Dimensión	Condición representada	Factor
Proyecto	Magnitud	Plazo de ejecución y recursos humanos	Importancia
	Complejidad	Conformación del equipo del proyecto	
Producto	Reutilización	Producto aislado o familia de productos	Trascendencia
	Ciclo de vida	Tiempo de vida (desarrollo y mantenimiento)	
	Magnitud	Tamaño y complejidad del sistema	Rango
	Criticidad	Tolerancia a fallas y a sus consecuencias	
Organización	Madurez	Nivel de madurez CMMI	Madurez
	Dependencia	Observancia de estándares	Dependencia

2.2 Selección de Métricas

Una vez definidos los factores, el paso siguiente fue seleccionar las métricas destinadas a cuantificar estos cinco factores adoptados y para ello se trató de encontrar un equilibrio entre indicadores, que sean suficientemente expresivos de las cualidades representadas y que a su vez sean de fácil evaluación y que puedan ser obtenidos de la industria. Además, los indicadores deben tener escalas de evaluación relacionados en forma directa con la trazabilidad, es decir que un valor mayor en la escala implique mayor viabilidad de ser trazado. Por último, para facilitar su interpretación se adoptó para todos un rango de lecturas uniforme comprendido entre uno y cinco.

Importancia del Proyecto, (i). El factor “Importancia” resume dos dimensiones características de los proyectos, como son su Magnitud y Complejidad. Se adopta como indicador de Magnitud el “esfuerzo de proyecto” (E), que es expresado en [personas-mes] y se determina a partir del modelo de predicción COCOMO II [5]. Para considerar la Complejidad del proyecto se hace referencia a la conformación del equipo de proyecto y se distinguen cuatro niveles, según sea: 1) completamente propio, 2) parcialmente externo con subcontratistas locales, 3) parcialmente externo con subcontratistas off-shore y 4) una combinación de los dos últimos. Luego, a partir de ambas dimensiones se obtiene el factor “Importancia” que se indica en la Tabla 2.

Tabla 2: Definición de la importancia del proyecto (*i*)

Complejidad (Conformación del equipo del proyecto)	Magnitud del Proyecto: <i>E</i> [personas-meses]		
	< 6	6 a 24	> 24
Completamente interno a la organización	1	2	3
Incluye subcontratistas locales	2	3	4
Incluye subcontratistas “off-shore”	3	4	5
Incluye subcontratistas locales y “off-shore”	4	5	5

Trascendencia del Producto, (t). El factor “Trascendencia” refleja la importancia que se asigna a la disponibilidad presente y futura de toda la documentación que debería ser aportada y operada a lo largo del ciclo de vida de un producto. Para definirlo se recurre a dos dimensiones del producto, que son el largo del ciclo de vida y la expectativa de reutilización y se lo hace a través de la Tabla 3. La primera dimensión contempla el largo del ciclo de vida, incluida la expectativa de su mantenimiento. La segunda está asociada a la potencialidad de reutilización de un producto, ya que esa circunstancia estimulará a realizar el mayor esfuerzo que demanda un proceso trazable detallado.

Tabla 3: Definición de la trascendencia de un producto (*t*)

Expectativas de reutilización	Ciclo de vida del producto [años]		
	< 1	1 a 5	> 5
Producto específico no repetible	1	2	3
Producto específico repetible	2	3	4
Núcleo base de una familia de productos	3	4	5

Rango del Producto, (r). El factor “Rango” es un indicador que resume tres dimensiones características del producto, como son su tamaño, complejidad y criticidad. Para reunir los dos primeros, es decir el tamaño y complejidad de un sistema, se recurre a los Puntos de Función ajustados tal como son definidos por el estándar ISO/IEC 20926 [6]. Por su parte, la criticidad es clasificada a partir de la valoración del riesgo según la tolerancia admitida a una contingencia del sistema [7]. Ambas dimensiones, puntos de función y valoración del riesgo definen el Rango mediante la Tabla 4.

Tabla 4: Definición del rango de un producto (*r*)

Valoración del riesgo según la tolerancia a una contingencia	Puntos de Función ajustados		
	< 100	100 a 1000	> 1000
Aceptable	1	2	3
Tan baja como sea razonable (ALARP) [8]	2	3	4
Intolerable	3	4	5

Madurez de la Organización, (m). El factor “Madurez” de una organización es clasificado según la escala adoptada por el modelo CMMI (Ver. 1.2) del SEI [9], que para la representación por etapas reconoce cinco niveles, que son: 1) Inicial, 2) Gestionado, 3) Definido, 4) Cuantitativamente gestionado y 5) Optimizado.

Dependencia de la Organización, (d). El factor “Dependencia” se refiere al nivel de autonomía que tiene una organización para operar según criterios propios o externos, y para ello se adopta la siguiente escala de clasificación: 1) absolutamente libre, 2) adhesión a estándares propios, 3) dependencia de estándares de clientes, 4) dependencia de estándares de casa matriz y 5) dependencia de estándares de clientes y casa matriz.

2.3 Ortogonalidad de Factores de Trazabilidad

De manera intuitiva, la ortogonalidad es un concepto matemático que generaliza la noción geométrica de perpendicularidad para el caso de espacios multidimensionales. Estrictamente hablando, dos direcciones son ortogonales si la proyección de una sobre la otra es nula, lo que equivale a decir que el producto escalar de sus componentes es cero.

En el campo de los sistemas de computación se ha adoptado el término “ortogonalidad” para significar independencia o disociación, es decir que varios factores son ortogonales entre sí cuando la variación que pueda experimentar alguno de ellos no afecta a los demás. En este caso se está haciendo referencia a factores hasta cierto punto independientes, que contribuyen a que un cierto proyecto sea trazable, y aquí debe reconocerse que esta ortogonalidad es una cualidad deseable pero que nunca es completamente absoluta.

3 Modelos de Predicción

Tal como fue anticipado, reconocer los factores que tienen mayor incidencia en la trazabilidad de los proyectos de software permite identificar y aprender a evitar las causas de las dificultades que se presentan con mayor frecuencia. El primer paso es entonces verificar que se han seleccionado los factores apropiados y para ello proponemos dos modelos dicotómicos que tienen como entradas los cinco factores ya descriptos y separan los casos estudiados en dos clases: “trazables” y “no trazables”.

3.1 Modelo Básico

En primera instancia se adoptó un modelo muy simple que discrimina entre clases “trazables” y “no trazables” observando si la totalidad de los “factores” de un cierto caso están por encima o por debajo de valores de frontera preestablecidos, mientras que una combinación de valores a ambos lados de la frontera conduce a casos especiales que son reconocidos como “atípicos”. Se trata de un análisis discriminante muy simple, compatible con la información relativamente escasa con la que se esperaba contar en la primera etapa del proyecto, pero que es suficiente para comprobar la correcta selección de los factores adoptados. Con este modelo queda definida una hipersuperficie discriminante a partir de la intersección de cinco hiperplanos, donde cada uno de ellos es normal al eje de un factor y lo corta a una distancia del origen igual a su valor de frontera.

Para poder visualizar la discriminación en forma gráfica se reducen las dimensiones del problema a tres, suponiendo que los factores son denominados X, Y, Z . En este ejemplo los valores de frontera son X_D, Y_D, Z_D y se considera un proyecto P_A con factores X_A, Y_A, Z_A . En la Fig. 1a se representa un sistema cartesiano ortogonal con los ejes X, Y, Z , los valores de frontera son representados por tres planos ortogonales que contienen el punto P_D y el proyecto queda indicado por el punto P_A . Se trata de un proyecto trazable ya que $X_A > X_D, Y_A > Y_D$ y $Z_A > Z_D$.

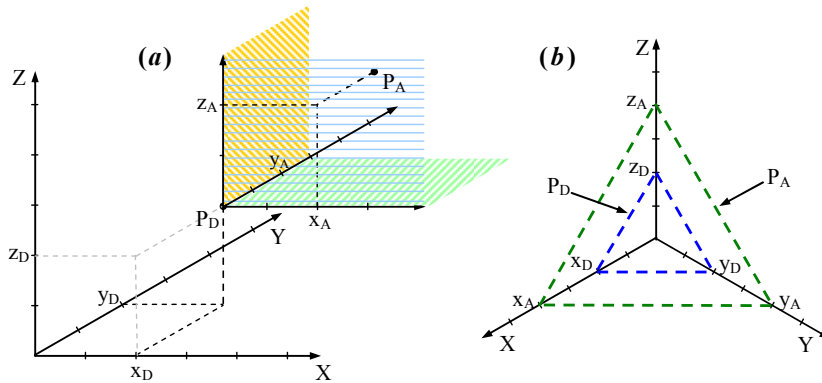


Fig. 1. Proyecto “P_A” y condiciones discriminantes “P_D” representados por (a) puntos en un hiperespacio y por (b) polígonos en un “diagrama de radar”.

Como puede observarse, aún para un ejemplo reducido a tres dimensiones, no resulta suficientemente clara la representación gráfica de un caso como el “P_A”, mucho menos clara sería la comparación de diferentes casos entre sí. Notar que es imposible visualizar el caso del modelo de trazabilidad completo con sus cinco dimensiones.

Para superar la dificultad de visualización, antes mencionada, recurrimos a un “diagrama radar”, que es un esquema plano con una cantidad de ejes radiales igual a la dimensión del modelo. En el ejemplo los ejes son tres, P_A y los valores discriminantes P_D quedan representados por puntos sobre cada uno de estos ejes X, Y, Z, como muestra la Fig. 1b. Como puede observarse, es posible definir un polígono discriminante que pase por los puntos X_D, Y_D, Z_D y separe el lugar geométrico de los proyectos trazables de los que no lo son. Por lo tanto, si el proyecto P_A es trazable quedará representado por otro polígono, exterior al anterior, que pasa por los puntos X_A, Y_A, Z_A.

Una vez adoptada una forma de representar este tipo de problemas, se regresa al modelo de trazabilidad de cinco dimensiones asociados a los factores [i, t, r, m, d] ya adoptados y se presentan diferentes casos en la Fig. 2.

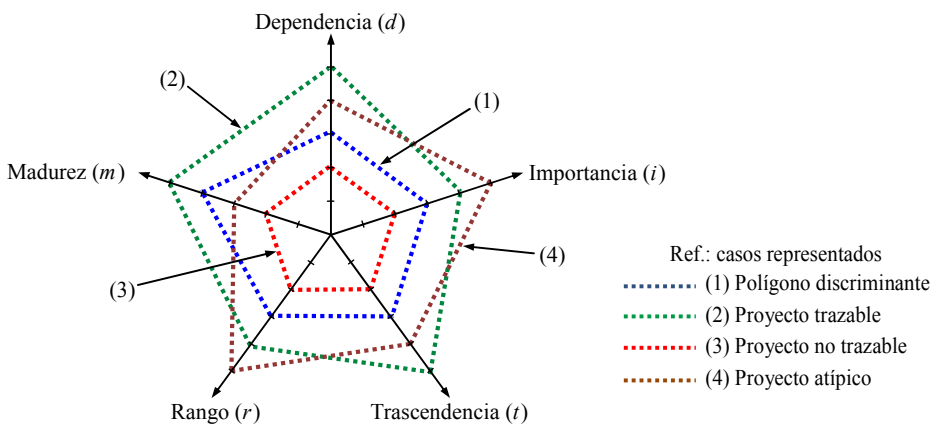


Fig. 2. Casos representados sobre el modelo de clasificación de cinco dimensiones

En la Fig. 2 se muestra: 1) el polígono discriminante, 2) un proyecto trazable, 3) un proyecto no trazable y 4) un proyecto atípico. En resumen, y tal como fue planteado, el “diagrama de radar” brinda dos tipos de resultados: *i*) clasificación de los proyectos en “trazables” y “no trazables” y *ii*) reconocimiento de los proyectos que se desvían significativamente de lo “normal” y que son calificados como “atípicos”. Estos últimos deben ser objeto de un análisis particular por admitir diversas interpretaciones, tales como: a) el caso no es contemplado por el modelo y éste debe ser perfeccionado para incluirlo, b) el caso está fuera del alcance de este tipo de modelo o c) los factores que definen a este caso atípico presentan inconsistencias o incongruencias y el modelo estaría advirtiendo una de estas circunstancias. Naturalmente, la existencia de un polígono discriminante capaz de clasificar al mayor número de casos en trazables y no trazables está supeditada a que se haya hecho una oportuna selección de los factores determinantes.

A continuación se muestra que el modelo propuesto y su medio de representación gráfica permite visualizar correctamente una diversidad de casos singulares, como son los siguientes ejemplos: 1) dos productos y proyectos diferentes a cargo de una misma empresa, 2) dos productos y proyectos iguales a cargo de diferentes empresas y 3) una empresa y proyecto con valores elevados en los factores (*i*), (*d*) y (*m*), a cargo del desarrollo de un producto con escaso rango (*r*) y poca trascendencia (*t*). Todos estos ejemplos son representados en la Fig. 3.

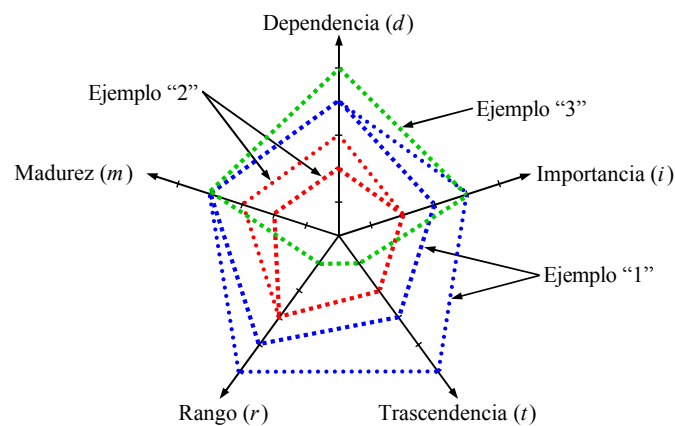


Fig. 3: Representación de ejemplos de casos singulares identificados como “1”, “2” y “3”

3.2 Modelo Avanzado

Como una evolución de la propuesta básica se presenta un modelo más preciso que se apoya en los mismos cinco factores pero que requiere de mayor cantidad y calidad de información para su definición. El objetivo es obtener un “índice de trazabilidad” ξ a partir de una expresión que tenga por argumentos a los cinco factores, que en forma general es expresada como:

$$\xi = f(i, t, r, m, d) ; \quad \xi = [0,1] \quad (1)$$

Se adopta como convención que la clase $\xi = 0$ agrupa a los casos donde por razones técnicas o económicas no es posible o conveniente aplicar un proceso de trazabilidad y la clase $\xi = 1$ contiene al grupo de los casos opuestos. Esto equivale a decir que la

discriminación entre ambas clases de casos queda establecida por una superficie compleja que es definida en el hiperespacio de cinco dimensiones, en lugar de los planos mutuamente ortogonales en el que se apoya el modelo básico. Con el fin de ilustrar lo dicho se reduce nuevamente el problema a uno de tres dimensiones X, Y, Z , mostrándose en la Fig. 4a un punto que representa cierto proyecto P_A y una superficie discriminante imaginaria que separa el lugar geométrico de los casos trazables y no trazables. Esta superficie reemplaza en el modelo avanzado a los planos del modelo básico ya mostrados en la Fig. 1a.

En la función discriminante propuesta se conocen sus argumentos pero obviamente se desconoce la expresión matemática que sea capaz de manifestar un comportamiento acorde a los datos provenientes de la industria con referencia al tema tratado. Más aún, esta función no solo debe ser establecida para un conjunto de casos, sino que debe ser ajustada en la medida que se disponga de mayor cantidad de información. Lo expuesto nos indujo a adoptar un modelo neuronal multicapa de perceptrones para representar la función desconocida, que seguramente es una expresión compleja, destinada a la evaluación del índice ξ . En este tipo de modelo los pesos sinápticos de la red y los umbrales de activación de cada unidad deben ser ajustados con una parte de los juegos de datos provenientes de casos relevados y el resto de estos casos son utilizados para su validación. A partir de lo expresado puede anticiparse que se trata de una red con cinco unidades de entrada y una unidad de salida, que es mostrada esquemáticamente en la Fig. 4b. Por el momento se desconocen la cantidad de capas de unidades ocultas necesarias y la composición de estas capas.

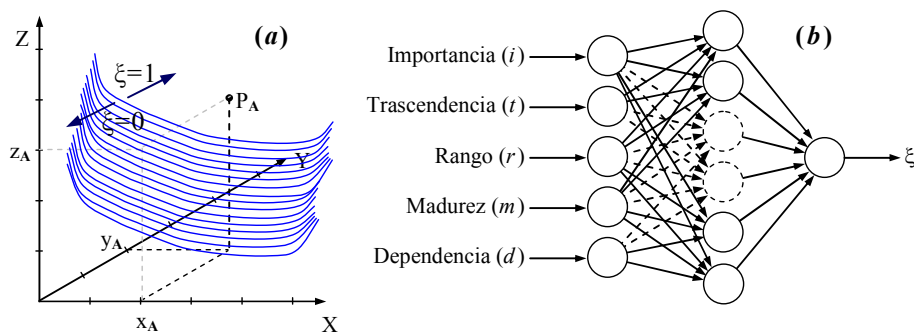


Fig. 4: (a) Representación de la hipersuperficie discriminante en un sistema de dimensiones reducidas y (b) esquema del modelo neuronal para la evaluación del índice de trazabilidad ξ

Puede anticiparse que, tan pronto se disponga de datos de la industria en la cantidad y calidad suficiente, será necesario perfeccionar el modelo a través de un proceso de validación, lo que dará lugar a los dos problemas siguientes: *i*) la necesidad de evaluar la conveniencia de utilizar otros factores y/o diferentes combinaciones de los mismos y *ii*) una vez identificado el modelo más apropiado, la necesidad de ajustar los valores de los umbrales de activación de manera de obtener su mejor desempeño en la clasificación. Por tratarse de un modelo de clasificación dicotómica, la herramienta natural para resolver los problemas planteados es el análisis ROC (Receiver Operation Characteristics) [9], de gran difusión en la actualidad. Este se apoya en la evaluación y representación (Curvas ROC) de la sensibilidad y especificidad calculadas a partir de los datos disponibles con el fin de identificar el mejor modelo para representar un cierto fenómeno.

4 Resultados Obtenidos

La obtención del tipo de información requerida por los modelos aquí propuestos es una tarea laboriosa, que tropieza con la natural reserva de las organizaciones a hacer públicos los detalles de sus procedimientos productivos y sus resultados. Esto se agudiza en el tema tratado, que como fue expuesto, muestra una amplia brecha entre lo efectivamente realizado y lo proclamado, lo que obedece a motivos estratégicos, comerciales y de imagen, entre otros. Para aliviar esta dificultad fue necesario poner énfasis sobre la naturaleza estrictamente estadística de las consultas realizadas.

Para concretar el relevamiento se diseñó una planilla con preguntas cuidadosamente formuladas, agrupadas según las tres entidades (producto, proyecto y organización) y destinadas a cuantificar las ocho dimensiones presentadas en la Tabla 1. Cada caso relevado tiene el alcance de un proyecto específico, se asienta en un formulario, reúne datos de las ocho dimensiones e incluye los resultados referidos a su trazabilidad.

Luego, a medida que se fueron realizando las entrevistas, los formularios debieron ser controlados para verificar su completitud y consistencia. Los resultados obtenidos hasta el momento muestran que, a pesar del empeño puesto por los encuestadores, aproximadamente la mitad de las entrevistas realizadas no condujeron a disponer de casos completos. Naturalmente esto las invalida para el objetivo planteado, pero aún así fueron de gran utilidad por contribuir a conocer mejor el problema y ayudar a confirmar la pertinencia de los factores considerados en los modelos. De las respuestas completas que se obtuvieron, algo más de la mitad fueron a su vez separadas por la presunción de que corresponden a casos atípicos y deberán ser estudiadas con mayor detalle una vez que el modelo básico se haya consolidado. En definitiva, sólo se pudo utilizar finalmente algo menos de la cuarta parte del total de consultas realizadas.

Al aplicar esos datos al modelo básico se comprobó que se polarizaban entre proyectos definitivamente trazables y no trazables, faltando casos intermedios que permitan definir con precisión el polígono discriminante que separa ambas clases. Por ello es necesario procurar mayor cantidad y variedad de información de campo, para lo cual se continúa trabajando en la realización de las encuestas.

Mientras se están realizando las actividades mencionadas, entrevistas y relevamientos, se trabaja paralelamente en perfeccionar desde un punto de vista computacional el segundo modelo, llamado avanzado. Ocurre que las redes neuronales ofrecen la ventaja de representar funciones desconocidas, inclusive altamente no lineales, pero deben ser ajustadas a partir de un proceso de prueba y error que en algunos casos es laborioso. Esto es debido a la ausencia de recomendaciones concluyentes sobre las arquitecturas de las redes de perceptrones que resultan más convenientes para cada tipo de modelo.

Para esta tarea se tomó como base el lote de los casos que se polarizaron con el modelo básico y se le incorporaron casos simulados que fueron reconstruidos a partir de la información disponible que no había sido utilizada por estar incompleta. Se dispuso así de un segundo lote de casos “imaginarios” destinados a la puesta a punto del modelo neuronal desde un punto de vista algorítmico y la validación de su capacidad de clasificación. Este trabajo implicó pruebas con varias redes, el ajuste de los parámetros de las mismas y la verificación de sus desempeños utilizando una estrategia de validación cruzada [9] a partir de la división del lote de datos en cinco subconjuntos ($k = 5$). La finalidad fue procurar independizar los resultados del modo en que se separan los datos destinados al ajuste del modelo y a su posterior validación.

Las pruebas realizadas permitieron llegar a las siguientes conclusiones con respecto

a la arquitectura del modelo neuronal propuesto: *i*) es suficiente una sola capa oculta, *ii*) se utilizaron funciones de activación sigmoideas, para lo cual fue necesario normalizar las entradas para que tomen valores en el rango $[0, 1]$, *iii*) una cantidad de entre siete y once unidades en la capa oculta brindaron los mejores resultados y *iv*) una vez ajustado el modelo demostró una excelente capacidad de clasificación.

Debe aclararse que las pruebas realizadas sirven para confirmar el potencial que ofrece el modelo denominado avanzado, que quedó disponible y deberá ser nuevamente ajustado una vez que se disponga de datos reales en la cantidad y variedad necesaria.

5. Conclusiones

Se desarrollaron dos modelos para estudiar las causas que obstaculizan la aplicación efectiva de la trazabilidad en la industria del software, que se apoyan en ciertos factores considerados determinantes y en métricas que permiten su cuantificación. Para el ajuste de estos modelos es necesario obtener información de la industria y para ello se está realizando un importante esfuerzo de trabajo de campo. Se utilizan dos modelos dicotómicos que discriminan los casos estudiados en “trazables” y “no trazables” utilizando cinco factores de entrada. En primera instancia se adoptó un modelo muy simple destinado a confirmar la conveniencia de estos factores de entrada y las métricas utilizadas, que es a su vez apto para trabajar con poca información. Posteriormente se desarrolló un modelo más avanzado basado en redes neuronales, que requiere más información y brinda una clasificación más precisa. Las evidencias han demostrado que la mayor dificultad está en la obtención de información de casos reales en la cantidad, calidad y variedad necesaria. Con este fin se continúa trabajando y se confía que a través de los modelos y criterios adoptados se alcanzarán los objetivos planteados.

Referencias

1. Gibbs, W.: Software's Chronic Crisis. Trends In Computing. Scientific American, vol. 86 (1994)
2. Gotel, O., Finkelstein, A.: Contribution Structures, Imperial College of Science, Technology and Medicine, Department of Computing, London (1995)
3. Rilling, J., Charland, P., Witte, R.: Traceability in Software Engineering - Past, Present and Future. CASCON 2007 Workshop Report, IBM Technical Report: TR-74-211 (2007)
4. Cleland-Huang, J.: Traceability Research: Taking the Next Steps. ICSE '11, Waikiki, Honolulu, HI, USA (2011)
5. Merlo-Schett, N.: COCOMO (Constructive Cost Model). Seminar on Software Cost Estimation WS 2002/2003, Department of Computer Science, University of Zurich, Switzerland (2003)
6. ISO/IEC 20926: IFPUG Functional Size Measurement Method. Software and Systems Engineering - Software Measurement, ISO/IEC International Standard, 2° Ed (2009)
7. Sommerville, I.: Ingeniería de Software. 7a Edición, Pearson Educación (2005)
8. Chrissis, M., Honrad, M., Shrum, S.: CMMI: Guía Para la Integración de Procesos y Mejora de Productos. 2a Edición, Pearson Educación (2009)
9. Orallo, J., Ramírez, Q., Ferri, C.: Introducción a la Minería de Datos. Pearson-Prentice Hall, 2a Edición, pp. 462--484 (2007)

Reconocimiento

A los Ings. Julio Massa (UNC) y Natalia Mira (IUA) por sus valiosos aportes y comentarios.