

Automatic Detection of Team Roles in Computer Supported Collaborative Work

P. García, J. M. Balmaceda, S. Schiaffino and A. Amandí

Abstract— Computer systems designed to support teamwork are environments that enable and enhance collaboration among users. These systems, named groupware, incorporate techniques that facilitate communication and coordination among team members to improve group performance. However, as not all team members are identical, it is important to study users' characteristics to build more productive teams.

Team Roles Theory allows obtaining the best possible performance taking into account individual skills, combining the weaknesses of each role with the strengths of other team members. Therefore, it is essential in group formation to consider mind the role each member is capable of playing. Currently, each team member has to complete lengthy questionnaires to determine his team role.

We propose an approach for automatic detection of team roles from the observation of user interactions in CSCW systems. Particularly, we apply Artificial Intelligence to obtain the team role that best suits a user's characteristics and behaviours. To achieve this, the performance of various classification algorithms was analyzed. The results show that our approach can detect team roles with high degrees of precision.

Keywords— Team Roles, Teamwork, Groupware, Computer Supported Collaborative Work.

I. INTRODUCCIÓN

EL TRABAJO en equipo a través de sistemas colaborativos es un tema que ha ganado una gran relevancia en los últimos años. Los avances en las tecnologías de redes y comunicación han propiciado la separación geográfica del personal de las organizaciones. Esto generó la necesidad de desarrollar nuevas herramientas que permitan la interacción de las personas para trabajar de forma distribuida y colaborativa. Por este motivo, surge una importante disciplina como es el Trabajo Colaborativo Soportado por Computadoras (CSCW, del inglés *Computer Supported Collaborative Work*) que estudia las características del trabajo en grupo y el desarrollo de herramientas colaborativas. Estas aplicaciones, que permiten la interacción entre usuarios para trabajar de forma conjunta y colaborativa, son denominadas aplicaciones para trabajo en grupo (groupware) o software colaborativo [3].

Mantener una colaboración efectiva entre los integrantes de un grupo es un problema difícil de resolver. La productividad

de los grupos estará determinada por la forma en que sus integrantes trabajan colaborativamente. En este contexto, las destrezas colaborativas y características de los integrantes de un equipo serán factores determinantes en el éxito del desarrollo de la actividad colaborativa. No todos los integrantes de un equipo son iguales; existen diferencias en sus comportamientos, conocimientos y habilidades para desarrollar una tarea. Estas diferencias pueden ser una fuente de conflictos y si no son manejadas correctamente causarán una importante disminución en la performance del equipo. Por esta razón, surgen las teorías de roles de equipo que estudian la manera de combinar personas con diferentes características para formar equipos de trabajo más productivos y eficientes.

Los roles de equipos son patrones de comportamiento característicos de la manera en la cual un miembro de un equipo interactúa con otros para facilitar el progreso del equipo [4]. Dentro de las teorías de roles de equipo pueden encontrarse varios modelos que definen diferentes roles y como deben ser combinados para maximizar la productividad de un grupo [2] [4] [16]. Este trabajo está basado en el modelo propuesto por Meredith Belbin [4] por ser una de las teorías de roles de equipos más difundida y validada [2]. Belbin propone en su modelo nueve roles de equipo: cerebro, investigador de recursos, coordinador, impulsor, monitor-evaluador, cohesionador, implementador, cohesionador y especialista. Junto con esta teoría de roles, Belbin ha desarrollado un instrumento, que consiste en una serie de cuestionarios, para detectar los roles que puede ocupar una persona.

En este trabajo se presenta una técnica de detección automática de roles de equipo utilizando diferentes técnicas de clasificación del área de Inteligencia Artificial. De esta manera, se logrará una nueva metodología de detección de roles que evita la necesidad de completar extensos cuestionarios y se realiza de forma transparente para los usuarios de las aplicaciones colaborativas. Este tipo de aplicaciones permiten registrar grandes cantidades de datos acerca de los procesos de interacción y la performance en la realización de las tareas de los integrantes de un grupo. Estos datos constituyen una importante fuente de información sobre la cual trabajar para caracterizar el comportamiento individual y grupal de un equipo de trabajo. Esta información nos permitirá obtener un modelo de usuario que describirá las preferencias y comportamientos de cada integrante de un grupo. Este modelo junto con la descripción de cada rol de equipo nos permitirá predecir los roles que mejor se adapten a una persona.

Este artículo está organizado de la siguiente manera. La

P. García, Universidad Nacional del Centro de la Provincia de Buenos Aires, Tandil, Buenos Aires, Argentina, pgarcia@exa.unicen.edu.ar.

J. M. Balmaceda, Universidad Nacional del Centro de la Provincia de Buenos Aires, Tandil, Buenos Aires, Argentina, jmbalma@exa.unicen.edu.ar.

S. Schiaffino, Universidad Nacional del Centro de la Provincia de Buenos Aires, Tandil, Buenos Aires, Argentina, sschia@exa.unicen.edu.ar.

A. Amandí, Universidad Nacional del Centro de la Provincia de Buenos Aires, Tandil, Buenos Aires, Argentina, amandi@exa.unicen.edu.ar.

Sección II describe en detalle el modelo de roles propuesto por Belbin. En la Sección III se presenta brevemente una revisión bibliográfica de los trabajos más destacados que aplican teorías de roles para mejorar la productividad de grupos virtuales. En la Sección IV se presenta la técnica propuesta para la detección automática de roles de equipos en sistemas para trabajo en grupo. En la Sección V se describen los experimentos realizados para validar la técnica propuesta. En la Sección VI se discute la precisión y validez de los modelos computacionales desarrollados para la detección de roles. Por último, en la Sección VII se exponen las conclusiones y trabajos futuros.

II. MARCO TEÓRICO

El concepto de rol ha sido definido como un conjunto de patrones de comportamiento esperados y atribuidos a alguien que ocupa una posición determinada en una unidad social [2] [5]. A partir del estudio de roles se han desarrollado un gran número de modelos y teorías que comienzan a estudiar el aporte de dichos roles en el trabajo en equipo. Estas investigaciones proponen los diferentes roles que las personas pueden ocupar dentro de un grupo de trabajo.

De esta manera, surge el concepto de roles de equipo. Belbin [4] fue el primer investigador en plantear una de las teorías de roles de equipo más difundidas en el área de formación de grupos. Con la aparición del concepto de roles de equipo en la década de los '80, se manifiesta una importante división en el concepto de roles entre roles funcionales y roles de equipo. Esta distinción entre estos dos tipos de roles es de gran importancia para que un grupo funcione efectivamente. Los roles funcionales abarcan las acciones físicas, actividades intelectuales, y objetivos específicos relacionados al trabajo de una persona. Por otro lado, los roles de equipo describen la naturaleza de las contribuciones que una persona realiza en beneficio del grupo. Estas contribuciones generalmente no son evidentes y no están asociadas a ningún requerimiento de trabajo. Los roles de equipo son adoptados naturalmente por las personas debido a su personalidad y comportamientos aprendidos, más que por capacidades técnicas y conocimientos específicos de una tarea.

En la literatura pueden encontrarse una diversa variedad de modelos de roles de equipos, además del modelo de Belbin, tales como Mumma [16], MTR-i (<http://www.profiles-r-us.com/mtri/>), Insights (<http://www.insights.com/index.aspx>), Thomas International (<http://www.thomasinternational.net/1/>), entre otras. Si bien la teoría de roles de equipo desarrollada por Belbin es la más difundida y validada [2], no existe ningún modelo de roles que abarque todas las situaciones posibles en la formación de equipos de trabajo. Cada modelo presenta diferentes puntos fuertes así como debilidades y limitaciones, por lo que la técnica presentada en las siguientes secciones para la detección de roles podría ser adaptada a cualquiera de los modelos de roles existentes.

Meredith Belbin [2] [4] [5] en sus investigaciones identificó nueve roles de equipo. A continuación se listan los

nueve roles de equipos describiendo la manera de comportarse, contribuir e interrelacionarse de las personas que pueden ocupar dichos roles.

Cerebro: es una persona creativa, poco ortodoxa e imaginativa. Además es muy efectivo para resolver problemas complejos dentro de un equipo pero tiene problemas para comunicarse y manejar un grupo de personas.

Investigador de recursos: es extrovertido, entusiasta y comunicativo. Explora oportunidades y desarrolla contactos. A pesar de ser un excelente negociador posee algunos problemas para mantener el interés en un proyecto una vez concluidas las primeras fases de desarrollo.

Coordinador: es maduro y confidente. Clarifica los objetivos del grupo, promueve la toma de decisiones y delega correctamente las tareas a realizar. El coordinador en ocasiones puede ser visto como un manipulador por su capacidad para organizar y delegar tareas a los miembros de un grupo.

Impulsor: es una persona desafiante, dinámica, sociable y que aplica presión al equipo. Posee la capacidad de superar cualquier obstáculo que se le presente. Las características negativas de las personas que ocupan este rol es que se enfadan fácilmente y provocan a otros integrantes del equipo generando conflictos.

Monitor-Evaluador: es una persona seria, estratega, juiciosa y crítica. Analiza todas las opciones y juzga las situaciones adecuadamente. Las personas con este perfil carecen de la habilidad de inspirar a otros miembros del equipo.

Cohesionador: es una persona tranquila, perceptiva, cooperativa y diplomática. Las personas que ocupan este rol saben escuchar a los miembros de su equipo, construyen consenso y evita cualquier tipo de problema dentro del grupo. Pero estas personas en muchas situaciones son indecisas e influenciables.

Implementador: es disciplinado, confiable, conservador y eficiente. A pesar de ser quien lleva las ideas a la práctica, es muy poco flexible y lento para responder a nuevas posibilidades e ideas.

Finalizador: es aplicado, meticoloso e impaciente. El finalizador busca errores y omisiones, entrega sus tareas en tiempo y forma, pero es propenso a preocuparse demasiado y prefiere no delegar actividades.

Especialista: es una persona muy dedicada capaz de centrarse en una meta y comenzar a trabajar por sí misma. Los especialistas proveen conocimiento pero solo contribuyen para completar una pequeña parte de una tarea o explicar en detalles ciertas técnicas a aplicar.

La mayoría de las teorías de roles de equipo han desarrollado diferentes herramientas para la detección de los roles que pueden ser asignados a una persona. Estas herramientas también pueden convertirse en importantes instrumentos de autoayuda para descubrir las habilidades actuales y potenciales de una persona. Particularmente, Belbin desarrolló el Inventario de Belbin de Auto-Percepción de Rol de Equipo (BTRSPI-9R, del inglés *Belbin Team-Role Self-*

Perception Inventory), un cuestionario compuesto de 7 secciones, cada una formada por 10 ítems que muestran diferentes situaciones de trabajo en equipo. Otra forma de asignar los roles de equipo de Belbin a los integrantes de un grupo es mediante el uso de un conjunto de ecuaciones para obtener los roles a partir de cuestionarios de personalidad como Cuestionario de Factores de Personalidad (16PF, del inglés *Sixteen Personality Factor Questionnaire*) [6] y el Cuestionario de Personalidad Ocupacional (OPQ, del inglés *Occupational Personality Questionnaire*) [21].

En este trabajo se plantea una solución alternativa a los cuestionarios de detección de roles utilizados actualmente para usuarios de sistemas colaborativos. Un aspecto importante de la técnica propuesta es que la detección se realiza de forma automática, y transparente a los usuarios. De esta manera, se evita a los usuarios contestar decenas de preguntas para conocer sus actitudes ante un grupo. En el uso de cuestionarios pueden ocurrir errores por una mala interpretación de las preguntas o si los cuestionarios son demasiado extensos los usuarios tienden a elegir respuestas arbitrariamente sin pensar cuidadosamente en ellas.

III. TRABAJOS RELACIONADOS EN DETECCIÓN DE ROLES

En los últimos años con el avance de las organizaciones virtuales y el continuo desarrollo de los sistemas para trabajo en grupo, los conceptos estudiados en el área de Psicología para la formación de equipos de trabajo han sido implementados exitosamente en el Trabajo Colaborativo Soportado por Computadoras. En la literatura se pueden encontrar artículos que muestran que el desempeño de grupos es mejor en aquellos donde existe un balance entre habilidades sociales y de resolución de tareas. Por ejemplo, en [22] se demuestra que existe una conexión entre los roles de equipos definidos y la necesidad de contar con grupos balanceados para obtener mejores desempeños en los equipos.

En lo que respecta a roles, en [13] se realiza un importante análisis de todas las teorías de roles que han sido desarrolladas a través de los años. Esta investigación apunta a comprender de forma más amplia los comportamientos de roles que los teóricos han definido. Además, en este trabajo intentan detectar roles a través del análisis cualitativo de una serie de videos que han sido filmados durante diversas secciones de trabajo en equipo.

El modelo de roles propuesto por Belbin [4] es una de las Teorías de Roles de mayor difusión y ha sido aplicada exitosamente en varios de los trabajos que estudian la formación de grupos. En [2] se reafirma la hipótesis del equilibrio de los roles, y el beneficio de utilizar los roles en la selección de miembros y composición de equipos. Continuando con esta hipótesis en [24] se propone un modelo de grupo balanceado que combina los roles de Belbin con conocimientos de los integrantes del grupo, el tipo de tarea a realizar y factores contextuales al grupo (localización, presiones de tiempo, entre otros). Otras investigaciones proponen modelados similares aplicados al área de ingeniería

de software para lograr equipos con una mayor performance y eficiencia [9] [10] [18] [24].

Otros trabajos [1] [10] tienen por objetivo lograr grupos balanceados según la teoría de combinación de roles desarrollada por Belbin. Para ello, en [10] se realiza una asignación dinámica de roles a equipos formados previamente teniendo en cuenta las diferencias culturales, de personalidad, y sus efectos sobre el trabajo en equipo. En cambio, en [1] se propone un agente inteligente que actúa como miembro de un equipo y toma los roles más adecuados en cada momento para balancear el grupo. Por otro lado, en [15] se propone una herramienta que genera grupos utilizando información referente a las personalidades, habilidades y debilidades de los miembros para evitar conflictos.

Como se puede observar en los trabajos enumerados anteriormente, existen un gran número de investigaciones que estudian la formación de equipos y trabajos que aplican estos conceptos a equipos virtuales. Sin embargo, no son muchos los trabajos que intentan detectar de manera automática los roles de equipo para hacer transparente esta actividad a los usuarios de las aplicaciones para trabajo en grupo. Algunos trabajos sobre la detección automática de roles de equipo son [7] [17] [23] [25]. En [7] [17] [23] estudian la influencia de los roles de grupo en el desempeño de los grupos de sistemas de aprendizaje basados en Web. Además, ambos trabajos proponen técnicas de minería de texto (*Text Mining*) para que los profesores puedan detectar los roles que cumplen sus alumnos. Por otro lado, en [25] se presenta una técnica de aprendizaje máquina para detectar automáticamente los roles funcionales de los participantes de reuniones presenciales. Esta investigación demuestra que es viable realizar dicha detección aunque sus resultados no son los suficientemente buenos como para aplicaciones reales.

A diferencia de los trabajos anteriores, en este trabajo se presenta una novedosa técnica para la detección automática de roles de equipo para usuarios de sistemas colaborativos. Esta técnica está basada principalmente en el análisis de las acciones que realicen las personas que trabajan de manera grupal y distribuida a través de aplicaciones para trabajo en grupo.

IV. DETECCIÓN DE ROLES UTILIZANDO INTELIGENCIA ARTIFICIAL

En esta sección se presenta la técnica propuesta para la detección automática de roles de equipo. En primera instancia, se presenta el esquema general del modelo propuesto. Luego, se describen en detalle las diferentes partes de la técnica propuesta.

A. Visión general de la técnica propuesta

Como se mencionó anteriormente, las aplicaciones para trabajo en grupo permiten la posibilidad de registrar grandes cantidades de datos acerca de los procesos de interacción y la performance en la realización de las tareas de los integrantes

de un grupo de trabajo. Estos datos constituirán el punto de partida para lograr la detección automática de roles de equipos en entornos colaborativos. En la Fig. 1 se resume el proceso propuesto para la detección de roles de equipo.

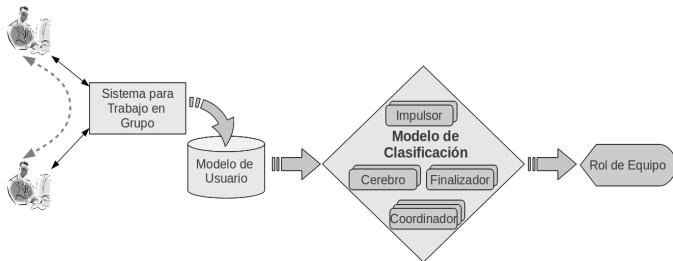


Figura 1. Esquema general para la detección automática de roles de equipo en sistemas para trabajo en grupo.

En primer lugar, se observará el comportamiento de los usuarios para almacenar sus interacciones dentro de sistemas colaborativos. Estas interacciones representan las acciones realizadas por un usuario dentro del sistema para alcanzar los objetivos de su equipo. Estas acciones registrarán tanto la utilización de las herramientas que brinde el entorno de colaboración como los aportes que realice el usuario en su grupo.

La información reunida acerca de cómo un usuario trabaja y se relaciona con los miembros de su equipo constituye el modelo de usuario. Estos datos serán utilizados por algoritmos de clasificación del área Inteligencia Artificial que permitan inferir el rol de un usuario. A continuación se explican cada unas de las partes del esquema presentado en la Fig. 1.

B. Interacciones observadas

En este trabajo se utilizó una aplicación para trabajo en grupo, desarrollada en la Universidad Nacional del Centro de la Provincia de Buenos Aires (UNCPBA), que permite capturar las interacciones de los usuarios trabajando colaborativamente para alcanzar un objetivo común. Dicha herramienta, al igual que la mayoría de los sistemas para trabajo en grupo, integra diversos componentes de colaboración como chat, foros de discusión, calendarios, mensajes instantáneos, entre otros; para brindar a los usuarios todos los mecanismos de colaboración, coordinación y comunicación necesarios para el trabajo en equipo.

La detección automática de roles de equipo se basa en el análisis de las acciones que un usuario realiza dentro del entorno para trabajo en grupo. Estas acciones están ligadas a habilidades colaborativas más que a preferencias individuales. Particularmente, se captura el desempeño de un usuario en el contexto de actividades grupales. Por ejemplo, serán almacenados los datos de participación en debates, aportes en la resolución de problemas, organización de actividades y recursos, apoyo a los miembros del grupo, entre otros. Esta información es de gran importancia para la detección de roles ya que las características de cada rol definidas en el modelo de Belbin pueden asociarse a acciones o comportamientos de las personas que integran un equipo de trabajo virtual. En la

Tabla I se resume la asociación que proponemos entre los roles de Belbin, sus características, y acciones observadas.

TABLA I. ACCIONES OBSERVADAS.

Roles de equipo	Posibles Acciones Observadas
Cerebro	Postea pocos mensajes Envía pocos mails Resuelve tareas
Investigador de Recursos	Gran número de contactos en la agenda de mails Muchos accesos al foro Muchos accesos al inicio del proyecto
Coordinador	Crea tópico de discusión Agrega novedades al grupo Asigna tareas
Impulsor	Muchos accesos al foro Mucha participación en las discusiones
Cohesionador	Mucha participación en los foros
Implementador	Realiza muchas tareas Termina las tareas en tiempo y forma
Finalizador	Entrega siempre a término sus tareas
Especialista	Provee material al grupo Agrega noticias

Como puede observarse en la Tabla I, no sólo se observan interacciones entre el usuario y la aplicación sino que también deben analizarse las interacciones entre los miembros de un grupo y los aportes realizados por cada usuario. La mayor parte de estas interacciones deben extraerse de canales de comunicación no estructurados como pueden ser los foros de discusión, chats, o mensajes instantáneos intercambiados entre los usuarios de la aplicación. Analizar estos textos no estructurados es un desafío computacional que requiere de un gran esfuerzo y en muchas situaciones no se obtienen los resultados deseados. Por esta razón, la técnica de detección de roles propuesta ha sido planteada inicialmente para ser aplicada sobre sistemas colaborativos estructurados. Es importante destacar que actualmente se están analizando y evaluando diferentes técnicas de minería de texto para que esta técnica pueda ser implementada sobre cualquier entorno de colaboración.

TABLA II. SENTENCIAS DE APERTURA.

Roles	Acciones	Sentencias de Apertura
Cerebro	Propone nuevas ideas y soluciones	“Intentemos...” “Yo creo que debemos intentar...” “Yo creo que ... porque ...”
Impulsor	Desafia al grupo	“No entiendo, ¿alguien puede...?” “¿Por qué...?”
	Discute	“Pero podría ocurrir que...”
	Muestra agresividad	“No estoy de acuerdo” “No estoy seguro”
Monitor - Evaluador	Analiza varias alternativas	“En vez de... probemos...” “En lugar de eso podríamos...” “Supongamos que...”
	Demora en la toma de decisiones	“Yo pienso que la mejor opción...”

Para lograr un sistema de colaboración estructurado se propone la utilización de un conjunto de sentencias de aperturas [19]. Dichas sentencias de aperturas son frases que deben ser seleccionadas antes de realizar un aporte a través de un chat o foro y brindan una visión general de la contribución

generada. De esta manera, los usuarios que deseen interactuar con sus compañeros a través de un foro de discusión, previamente deberán seleccionar la sentencia de apertura que mejor se adapte a su colaboración. Esta estrategia permitirá entender en tiempo de ejecución los aportes realizados por los usuarios. En la Tabla II se presenta, a modo de ejemplo, un conjunto de posibles oraciones de aperturas y su relación con algunas de las acciones que permiten la detección de roles.

C. Técnicas de clasificación aplicadas

La clasificación es una técnica de aprendizaje automático supervisado que permite a partir de instancias de datos construir hipótesis generales que a su vez son utilizadas para hacer predicciones sobre instancias de datos futuros [14]. Para alcanzar este objetivo se crea un modelo basado en la distribución de los datos que luego es usado para clasificar nuevos datos. Dado el modelo de clasificación, se puede predecir la clase de un nuevo conjunto de datos dado como entrada. En este trabajo se propone la construcción de un clasificador que provea una función que dado un registro de datos, como las acciones observadas, le asigna una de las distintas clases predefinidas, es decir los roles de equipos de Belbin.

El proceso de construcción del modelo de clasificación se compone de dos etapas:

Construcción del modelo (Aprendizaje): se construye un conjunto de datos de entrenamiento. Este conjunto de datos son tuplas donde cada una representa las acciones de un usuario y la clase a la que pertenece.

Evaluación del modelo: se estima la exactitud del modelo basándose en un conjunto de evaluación, comparando la etiqueta conocida de una tupla de evaluación con el resultado de aplicar el modelo de clasificación.

Las técnicas que serán analizadas para la detección de roles de equipos son: clasificador bayesiano, arboles de decisión y reglas de asociación. En las siguientes subsecciones se describen brevemente cada una de las técnicas de clasificación que se utilizaron en la fase de experimentación y evaluación de este trabajo.

1) Clasificador Bayesiano

Un clasificador bayesiano, como por ejemplo el algoritmo Naive Bayes, es un clasificador probabilístico simple basado en la aplicación del Teorema de Bayes con una fuerte suposición de independencia. Es decir, un clasificador que asume que la presencia de una característica particular de una clase no está relacionada con la presencia de ninguna otra característica. En [26] se intenta demostrar las razones teóricas de la eficacia de los clasificadores bayesianos en complejas situaciones del mundo real. Este clasificador es muy eficiente en diferentes dominios, debido a que es muy robusto ante atributos irrelevantes. Además, requiere de pequeños conjuntos de datos de entrenamiento para poder estimar los parámetros necesarios para la clasificación.

La Fig. 2 muestra la representación gráfica del clasificador Naive Bayes para detección de roles de equipos. Este gráfico captura la suposición fundamental de Naive Bayes, donde todos los atributos observados sobre el sistemas para trabajo

en grupo son independientes entre sí conocida la clase (en este caso los roles de equipos), por lo que no existen arcos entre los atributos y existe un arco del nodo raíz a cada uno de los atributos.



Figura 2. Modelo de clasificación del algoritmo Naive Bayes.

2) Árboles de Decisión

Un árbol de decisión es una estructura en la cual los datos son divididos de acuerdo a un criterio (test). Cada nodo del árbol denota un test sobre un atributo. Cada rama representa un resultado del test, y las hojas del nodo representan clases o distribución de clases [11]. Cada instancia de datos tiene varios atributos, uno de los cuales (el objetivo o atributo clase), indica la clase a la cual pertenece cada instancia. Las principales ventajas de este tipo de clasificador es que es muy sencilla su interpretación y es posible convertir el árbol en un conjunto de reglas para clasificar la entrada. Pero a diferencia de otros clasificadores el tiempo de entrenamiento es muy elevado, no maneja correctamente variables continuas y un error en una rama desencadena errores en las ramas inferiores. Un algoritmo de este estilo puede ser el J48 (basado en el conocido C4.5 [20]) que es un algoritmo desarrollado por los creadores de la herramienta Weka (<http://www.cs.waikato.ac.nz/ml/weka/>), la cual provee una colección de algoritmos de aprendizaje automático para tareas de minería de datos.

En la Fig. 3 se muestra parte del árbol de decisión generado para la detección de roles de equipos. Como puede observarse en dicha figura las hojas corresponden a los diferentes roles de equipos y los nodos son atributos que de acuerdo a sus valores permiten clasificar los datos de entrada en un rol en particular.

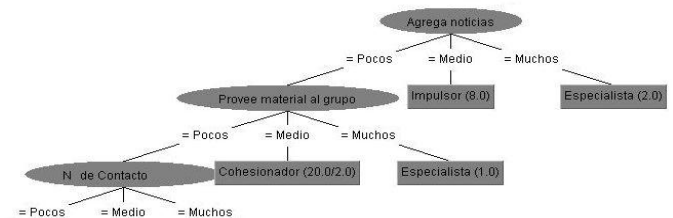


Figura 3. Modelo de Clasificación del algoritmo J48.

3) Clasificación basada en reglas

RIPPER [8] es un algoritmo de extracción de reglas para clasificación desde un conjunto de datos. Este algoritmo es una versión optimizada de IREP [12]. Estos algoritmos para aprendizaje de reglas trabajan en dos etapas. En la primera fase inducen un conjunto de reglas generales sobre los datos de entrenamiento. Luego, se mejora dicho conjunto de reglas usando una etapa de optimización que poda o ajusta ciertas

reglas ya que el conjunto inicial suele ser excesivamente grande y redundante. Este proceso de optimización hace que este algoritmo tienda a ser bastante lento en la generación del modelo de clasificación. A continuación se muestra el conjunto de reglas obtenidos a partir del algoritmo RIPPER para la clasificación de roles de equipo.

1. (Número de Contacto = Muchos) and (Agrega novedades al grupo = Pocos) => Rol=Investigador-Recursos
2. (Provee material al grupo = Muchos) => Rol=Especialista
3. (Termina las tareas en tiempo y forma = Medio) and (Propone nuevas ideas y soluciones = Si) => Rol=Cerebro
4. (Termina las tareas en tiempo y forma = Muchos) => Rol=Implementador
5. (Agrega novedades al grupo = Muchos) and (Agrega noticias = Medio) => Rol=Impulsor
6. (Agrega novedades al grupo = Pocos) and (Demora en la toma de decisiones = Si) => Rol=Monitor-Evaluador
7. (Provee material al grupo = Medio) => Rol=Cohesionador
8. (Número de Contacto = Pocos) => Rol=Cohesionador
9. => Rol=Coordinador (26.0/2.0)

V. EVALUACIÓN EXPERIMENTAL

En esta sección se presentan los resultados de los experimentos realizados para evaluar el modelo de detección de roles propuesto y el desempeño de las distintas técnicas de clasificación que fueron aplicadas. En primer lugar, se describen los datos utilizados en las pruebas experimentales. Luego, se presentan los resultados obtenidos en la detección de roles aplicando clasificadores bayesianos, árboles de decisión y reglas de asociación.

A. Datos experimentales

Los datos utilizados en las pruebas realizadas fueron recolectados de la interacción de 94 estudiantes de cuarto año de la carrera de Ingeniería de Sistemas de la UNCPBA. Estos alumnos, durante el segundo cuatrimestre del 2009, realizaron un curso sobre Inteligencia Artificial en cual utilizaron nuestro sistema para trabajo en grupo con el objetivo de resolver de manera colaborativa una serie de ejercicios prácticos. Los datos se utilizaron para analizar la precisión de las diferentes técnicas de clasificación propuestas para la detección de roles de equipos.

En primer lugar, fue necesario discretizar cada uno de los atributos presentes en el modelo de usuario debido a que algunos algoritmos de clasificación usados no aceptan datos continuos como entrada. Básicamente se realizó una discretización en tres rangos (*Mucho*, *Medio*, *Poco*) para aquellos valores continuos como por ejemplo participación en foros, tareas realizadas, e-mails enviados, entre otros. Los rangos utilizados fueron asignados dependiendo de la distribución de los datos obtenidos en cada atributo. Por ejemplo, para el atributo "*E-Mails Enviados*" se definieron los siguientes valores para cada rango:

- *Pocos*: [0;10]
- *Medio*: (10; 30]
- *Muchos*: (30; infinito)

Luego, se etiquetaron los datos. Las etiquetas o clases utilizadas fueron los nueve roles de equipos propuestos por Belbin. Para llevar a cabo este proceso de etiquetado de datos los profesores del cursos, con amplios conocimientos sobre el modelo de roles de equipo utilizado, especificaron los roles de cada estudiantes según el desempeño que habían exhibido a través de diferentes trabajos presenciales. De esta manera, logramos un conjunto de datos donde se le incorporó el rol de alumno como un nuevo atributo en los datos de entrada que permitió entrenar y evaluar los diferentes modelos de clasificación.

B. Pruebas realizadas

Las pruebas realizadas se basan en la creación de diferentes modelos de clasificación con el conjunto de datos obtenido. Luego, se analizaron distintas métricas para decidir cuál es la técnica que mejor se adapta al proceso propuesto para la detección automática de roles. Particularmente, en este trabajo se analizó la matriz de confusión, la precisión y la cobertura (*recall*). A continuación se describen cada una de estas métricas para un mejor entendimiento de los resultados arrojados por los diferentes algoritmos de clasificación:

Matriz de confusión: es una matriz cuadrada de $N \times N$, donde N es la cantidad de clases del clasificador. Las columnas representan las clases en las que se clasificaron los ítems, y las filas corresponden a las clases a las que pertenecen dichos ítems. Un clasificador ideal posee en esta matriz valores distintos de ceros en la diagonal principal, y cero en el resto de las celdas.

Precisión: la precisión para una clase A está definida como el número de verdaderos positivos (el número de ítems que fueron clasificados en A y afirmativamente pertenecen a la clase A) dividido por el total de ítems clasificados en A . En otras palabras, un valor de precisión de 1.0 (valor máximo) para una clase A significa que todos los ítems clasificados como pertenecientes a esta clase, verdaderamente pertenecen a A .

Cobertura: la cobertura para una clase A está definido como el número de verdaderos positivos dividido por el total de elementos que verdaderamente pertenecen a la clase A . Un valor de cobertura igual a 1.0 para una clase A significa que todos los ítems que pertenecen a la clase A van a ser clasificados en A .

En las siguientes subsecciones se presentan los resultados obtenidos con cada uno de los algoritmos de clasificación aplicados.

TABLA III. RESULTADOS DE LAS TÉCNICAS DE CLASIFICACIÓN.

Técnica Roles	Naive Bayes		J48		Ripper	
	Precisión	Cobertura	Precisión	Cobertura	Precisión	Cobertura
Cerebro	1	0,857	0,778	1	0,75	0,857
Investigador de Recursos	0,5	0,2	0,75	0,6	0,833	1
Coordinador	1	1	1	0,958	0,852	0,958
Impulsor	1	0,889	1	0,889	0,889	0,889
Monitor-Eval.	0,471	0,727	0,6	0,818	0,786	1
Cohesionador	0,826	0,905	0,857	0,857	1	0,714
Implementador	1	0,778	1	1	1	1
Finalizador	0	0	0	0	0	0
Especialista	1	1	1	0,833	1	1
Resultados General	0.851	0.84	0.87	0.872	0.878	0.883

1) Naive Bayes (Clasificador Bayesiano)

En la Fig. 4 se puede observar la matriz de confusión del algoritmo Naive Bayes y en la Tabla III se muestran las medidas obtenidas con este clasificador. Es importante destacar la exactitud que tiene Naive Bayes para clasificar los roles *Coordinador* y *Especialista*. Sin embargo, para *Investigador de Recursos* o *Monitor-Evaluador*, sus medidas son relativamente bajas. En el caso especial del rol *Finalizador*, se puede ver que en todas las pruebas realizadas, no ha sido clasificado correctamente ningún ítem perteneciente a esta clase.

```

a b c d e f g h i <-- classified as
6 0 0 0 1 0 0 0 0 | a = Cerebro
0 1 0 0 3 0 0 1 0 | b = Investigador-Recursos
0 0 24 0 0 0 0 0 0 | c = Coordinador
0 1 0 8 0 0 0 0 0 | d = Impulsor
0 0 0 0 8 3 0 0 0 | e = Monitor-Evaluador
0 0 0 0 2 19 0 0 0 | f = Cohesionador
0 0 0 0 2 0 7 0 0 | g = Implementador
0 0 0 0 1 1 0 0 0 | h = Finalizador
0 0 0 0 0 0 0 0 6 | i = Especialista
    
```

Figura 4. Matriz de Confusión de Naive Bayes.

2) J48 (Árboles de Decisión)

En la Fig. 5 se puede observar la matriz de confusión del algoritmo J48 y en la Tabla III se presentan los resultados obtenidos con este clasificador. Como puede observarse se obtuvieron excelentes resultados en la clasificación del rol *Implementador*. En particular todos los roles han logrado valores superiores a 0.6 de precisión y cobertura, excepto el rol *Finalizador* que al igual que con el algoritmo Naive Bayes ninguna tupla perteneciente a esta clase ha sido clasificada correctamente.

3) Ripper (Clasificación basada en reglas)

En la Fig. 6 se puede observar la matriz de confusión del algoritmo Ripper y en la Tabla III se presentan los resultados de precisión y cobertura arrojados por este clasificador basado en reglas. Analizando la sección de resultados del algoritmo Ripper en la Tabla III puede observarse que este algoritmo logró clasificar todas las clases con una precisión superior a 0.75, excepto el rol *Finalizador* el cual arroja una precisión de 0 en todas las pruebas realizada.

```

a b c d e f g h i <-- classified as
7 0 0 0 0 0 0 0 0 | a = Cerebro
0 3 0 0 2 0 0 0 0 | b = Investigador-Recursos
1 0 23 0 0 0 0 0 0 | c = Coordinador
0 1 0 8 0 0 0 0 0 | d = Impulsor
0 0 0 0 9 2 0 0 0 | e = Monitor-Evaluador
0 0 0 0 3 18 0 0 0 | f = Cohesionador
0 0 0 0 0 0 9 0 0 | g = Implementador
1 0 0 0 1 0 0 0 0 | h = Finalizador
0 0 0 0 0 1 0 0 5 | i = Especialista
    
```

Figura 5. Matriz de Confusión de J48.

```

a b c d e f g h i <-- classified as
6 1 0 0 0 0 0 0 0 | a = Cerebro
0 5 0 0 0 0 0 0 0 | b = Investigador-Recursos
1 0 23 0 0 0 0 0 0 | c = Coordinador
0 0 1 8 0 0 0 0 0 | d = Impulsor
0 0 0 0 11 0 0 0 0 | e = Monitor-Evaluador
0 0 2 1 3 15 0 0 0 | f = Cohesionador
0 0 0 0 0 0 9 0 0 | g = Implementador
1 0 1 0 0 0 0 0 0 | h = Finalizador
0 0 0 0 0 0 0 0 6 | i = Especialista
    
```

Figura 6. Matriz de Confusión de Ripper.

VI. ANÁLISIS DE RESULTADOS

Los resultados logrados en este estudio son muy alentadores. Analizando la Tabla III, puede observarse que el desempeño general de las tres técnicas de clasificación utilizadas es muy similar en cuanto a los valores de precisión y cobertura. Sin embargo, realizando un estudio más detallado sobre cada una de las clases podemos observar que el algoritmo Naive Bayes, a pesar de tener la precisión promedio más baja, es el clasificador que más roles detecta con la máxima precisión (ver Fig. 7). En cambio, el algoritmo Ripper permite obtener un modelo de clasificación de mayor rendimiento global en la precisión de detección de roles pero, como se muestra en la Fig. 7.

En la Fig. 8 se presentan los resultados de cobertura de las técnicas de clasificación estudiadas. A simple vista puede observarse que los mejores resultados de cobertura son arrojados por el algoritmo Ripper. Los usuarios con roles de *Investigador de Recursos*, *Monitor-Evaluador*, *Implementador* y *Especialista* fueron correctamente.

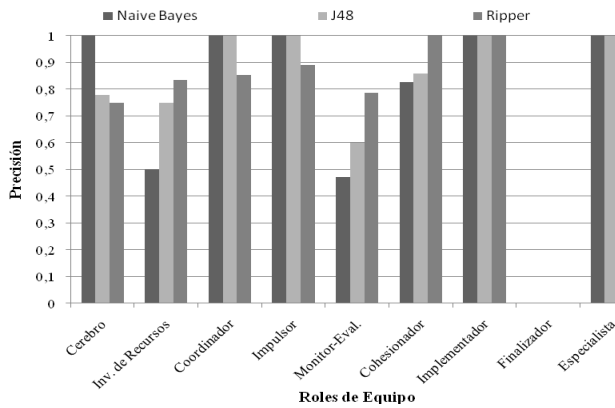


Figura 7. Precisión de las técnicas de clasificación.

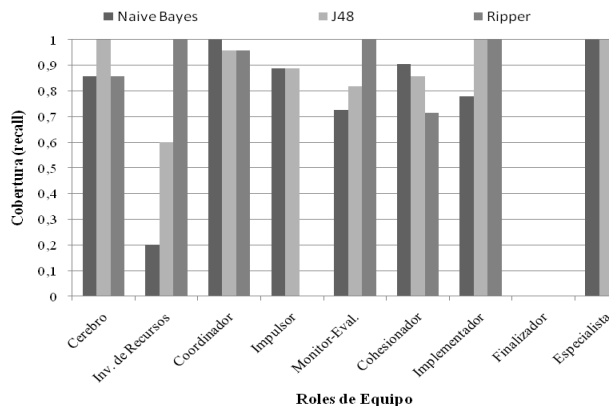


Figura 8. Cobertura de las técnicas de clasificación.

El principal problema de este estudio se encuentra en la predicción del rol de *Finalizador*, el cual no pudo ser detectado con ninguno de los algoritmos de clasificación. En este caso, el problema surge por el conjunto de datos utilizado para el armado del modelo de clasificación. En estos datos sólo dos alumnos mostraron cumplir el rol de *Finalizador* lo que hace que sea imposible aprender desde los datos las características que definirán a este rol de equipo.

Otra de las limitaciones importante de los clasificadores para la detección de roles de equipo es que encapsulan a los usuarios en un único rol. Este es un aspecto importante a tener en cuenta ya que los modelos de roles de equipo establecen que una persona puede desempeñar naturalmente más de un rol. En este contexto, los clasificadores probabilísticos como Naive Bayes son los más adecuados ya que permiten conocer la probabilidad de los datos de entrada de pertenecer a las diferentes clases del modelo de clasificación. De esta manera, es posible asociar varios roles a un único usuario dependiendo de las probabilidades obtenidas para cada clase.

VII. CONCLUSIONES Y TRABAJOS FUTUROS

En este trabajo se presentó un modelo para la detección automática de roles de equipo para usuarios de sistemas para trabajo colaborativo. Para alcanzar este objetivo, se aplicaron distintas técnicas de clasificación del área de Inteligencia Artificial sobre los datos obtenidos de la observación las

acciones realizadas por los usuarios dentro de sistemas para trabajo en grupo. Los resultados demuestran una importante precisión del modelo propuesto que avalan la utilización de esta técnica, como alternativa a los cuestionarios de detección de roles, para la formación de equipos de trabajo más productivos y eficientes.

Actualmente, estamos en el proceso de diseño de técnicas basadas en minería de textos para detectar las intenciones de los usuarios en las contribuciones que realicen al grupo. De esta manera, analizando los textos generados por los usuarios, se eliminará el uso de sentencias de aperturas y la técnica de detección de roles podría ser aplicada en sistemas colaborativos no estructurados.

Por otro lado, en los próximos trabajos se planea realizar nuevos experimentos con un mayor número de usuarios que permitan generar un conjunto de datos más extenso. También se pretende comparar los resultados obtenidos por la técnica propuesta con los resultados arrojados por el cuestionario de detección de roles desarrollado por Belbin. Esto brindaría una forma diferente de validar la técnica presentada en este artículo.

AGRADECIMIENTOS

Este trabajo fue parcialmente financiado por ANPCyT (Argentina) a través del Proyecto PICT 2011 N° 0366.

REFERENCIAS

- [1] R. A. Aguilar, A. Antonio y R. Imbert, "Searching Pancho's Soul: An Intelligent Virtual Agent for Human Teams". IEEE Computer Society, pp. 568—571. (2007).
- [2] A. Aritzeta, S. Swales y B. Senior, "Belbin's Team Role Model: Development, Validity and Applications for Team Building". Journal of Management Studies, pp. 96-118. (2007).
- [3] L. J. Bannon y K. Schmidt, "CSCW: four characters in search of a context". North-Holland Publishing Co., pp. 3—16. (1991).
- [4] R. M. Belbin, "Management teams: Way they succeed or fail". Oxford: Butterworth-Heinemann. (1981).
- [5] R. M. Belbin, "Team roles at work". Oxford: Butterworth-Heinemann. (1993).
- [6] R. B. Cattell, H. W. Eber y M. M. Tatsuoka, "The handbook for the Sixteen Personality Factor Questionnaire". Champaign, IL: Institute for Personality and Ability Testing. (1970).
- [7] G. Chen, C. Wang, K. Ou y B. Liu, "Using role theory in monitoring web group learning systems". Computers in Education, 2002 International Conference on Computers in Education (ICCE'02), p. 884. (2002).
- [8] W. W. Cohen. "Fast effective rule induction". In ICML, page 115-123. (1995).
- [9] F. Q. B. Da Silva. y A. C. F. César, "An Experimental Research on the Relationships between Preferences for Technical Activities and Behavioural Profile in Software Development". SBES '09. XXIII Brazilian Symposium, pp. 126 - 135, Fortaleza, Brasil. (2009).
- [10] G. A. Dafoulas y L. A. Macaulay, "Facilitating Group Formation and Role Allocation in Software Engineering Groups". IEEE Computer Society, pp. 352. (2001).
- [11] Q. Ding, Q. Ding, y W. Perrizo, "Decision tree classification of spatial data streams using Peano Count Trees", ACM 2002, pp. 413-417. (2002).
- [12] J. Fürnkranz y G. Widmer. "Incremental reduced error pruning". In ICML, page 70-77. (1994).
- [13] J. A. R. Guasch y C. M. Gonzalez, "Análisis de roles de trabajo en equipo: un enfoque centrado en comportamientos". Universidad Autónoma de Barcelona. (2006).
- [14] S. B. Kotsiantis. "Supervised Machine Learning: A Review of Classification Techniques". Informatica, Vol. 31, pp. 249-268. (2007).

- [15] S. Licorish, A. Philpott, y S. G. MacDonell, "Supporting agile team composition: A prototype tool for identifying personality (In)compatibilities". In Proceedings of the 2009 ICSE Workshop on Cooperative and Human Aspects on Software Engineering (CHASE '09). IEEE Computer Society, Washington, DC, USA, 66-73. (2009).
- [16] F. Mumma, "Team-Work & Team-Roles". 3rd Ed., King of Prussia: HRDQ, (1992).
- [17] K. Ou, C. Wang y G. Chen, "Identify group roles by text mining on group discussion in a Web-based learning system". Machine Learning and Cybernetics. Proceedings of 2005 International Conference, Vol. 9, 18-21. (2005).
- [18] V. Pieterse, D. G. Kourie, y I. P. Sonnekus, "Software engineering team diversity and performance". South African Institute for Computer Scientists and Information Technologists, pp. 180—186. (2006).
- [19] W. Pohl, A. Kobsa y O. Kutter, "User Model Acquisition Heuristics Based On Dialogue Acts", International Workshop on the Design of Cooperative Systems, pp. 471—486. (1995).
- [20] J. R. Quinlan. "C4.5: Programs for Machine Learning". Morgan Kaufmann Publishers Inc., San Francisco, CA, USA. (1993).
- [21] P. Saville, R. Holdsworth, G. Nyfield, L. Cramp y W. Mabey, "Occupational Personality Questionnaire Manual". Esher: SHL. (1992).
- [22] B. Senior, "Team roles and team performance: Is there 'really' a link?". Journal of Occupational and Organizational Psychology. (1997).
- [23] C. Wang, K. Ou, C. Liu, B. Liu y G. Chen. "Using a Text Miner to Identify Group Roles of Students in a Web-based Learning System". International Journal of Learning, Volume 9. (2002).
- [24] M. Winter, "Developing a group model for student software engineering teams". University of Saskatchewan. (2004).
- [25] M. Zancanaro, B. Lepri, y F. Pianesi, "Automatic detection of group functional roles in face to face interactions". In Proceedings of the 8th international conference on Multimodal interfaces (ICMI '06). (2006).
- [26] H. Zhang, "The Optimality of Naive Bayes". AAAI Press (2004).



Patricio García received the his degree in Systems Engineering from Universidad Nacional del Centro de la Provincia de Buenos Aires (UNCPBA), Tandil, Argentina, in 2006. He also received his Master degree in Systems Engineering in 2008 from the same institution. His research interests are Artificial Intelligence, intelligent agents, user profiling and computer supported collaborative systems.



José María Balmaceda received his bachelor degree in Systems Engineering from Universidad Nacional del Centro de la Provincia de Buenos Aires (UNCPBA), Tandil, Argentina, in 2010. Currently, he is PhD student in the same Institution and he holds a PhD CONICET scholarship. He is working under the supervision of Silvia Schiaffino. His research interests are artificial intelligence, user profiling and text analysis.



Silvia Schiaffino is Adjunct Researcher at CONICET (Argentina) and full Adjunct Professor at Universidad Nacional del Centro de la Provincia de Buenos Aires (UNCPBA), in Tandil, Argentina. She is also member of ISISTAN Research Institute, Tandil. She obtained her Master degree in Systems Engineering in 2001 at UNCPBA, and her PhD in Computer Science in 2004 at the same university. Her current research interests include user profiling, agents, and recommender systems, in different application domains.



Analia Amandi is full Professor at the Computer Science Department of Universidad Nacional del Centro de la Provincia de Buenos Aires (UNCPBA, Tandil, Argentina) and a Principal Researcher at CONICET. She is also Director of the Intelligent Agents Group at ISISTAN Research Institute, Tandil. She obtained her Bachelor degree in 1990 at Universidad Nacional de La Plata, and her Phd in Computer Science in 1997 at UFRGS, Porto Alegre, Brasil. Her main research interests are user profiling, intelligent agents and knowledge management.