Prevención de incidentes informáticos mediante aplicación de técnicas de Minería de Datos.

Corso, Cynthia Lorena
García, Mario Alejandro
Gibellini, Fabián
Ciceri, Leonardo Ramón
Romero, Fernando
Universidad Tecnológica Nacional, Facultad Regional Córdoba

Abstract

En este trabajo se exponen los resultados parciales correspondiente al primer año de ejecución del proyecto de investigación "Generación de Modelo Descriptivo para la prevención de incidentes de equipos informáticos en el contexto de laboratorio de sistemas", homologado por la Secretaria de Ciencia y Técnica de la Universidad Tecnológica Nacional. Esta investigación es de carácter descriptivo, ya que se intenta explicar y caracterizar los aspectos que propician la presentación de incidentes.
El objetivo central es el descubrimiento de factores que tienen mayor grado de influencia en la presentación de incidentes informáticos.
Se propone como alternativa de solución la utilización de métodos estadísticos incluidos en metodologías como la Ingeniería en Confiabilidad pilar del Mantenimiento que facilita el diagnostico de incidentes, complementada con el uso de métodos pertenecientes a la rama de la Minería de Datos.
La muestra de registros de incidentes considerados en esta investigación corresponde al periodo 2004 al 2012.
El modelo de investigación adoptado es mixto, es decir combina el enfoque cuantitativo y cualitativo.

Palabras Claves
Incidente Informático, Laboratorio Informático, Mantenimiento, Minería de Datos, Ingeniería de Confiabilidad, Predicción, Algoritmos de Asociación.

Introducción

En la actualidad las organizaciones están insertas en un contexto de cambios continuos, por ello las empresas privadas tanto como las públicas deben tener la capacidad de adaptarse, aprender cómo resolver problemas y generar conocimiento, para establecer nuevos métodos en pro de la resolución de los mismos.[1]
Es por ello que las organizaciones invierten tiempo y costos en el diseño de estrategias y análisis de riesgos para poder analizar las fallas, que puedan surgir, con el propósito de lograr mantener en el tiempo el funcionamiento normal de máquinas y equipos que generan un producto o servicio. [2]
El uso de las computadoras en el ámbito educativo, ha sido un proceso en desarrollo desde 1960, caracterizado en sus inicios por los proyectos estatales a gran escala. Sin embargo es indiscutible, para muchos, que desde la década de 1990 se produjo un avance significativo en la necesidad de dotar a las escuelas, institutos y universidades con computadoras. [3]
En este último caso muchas universidades cuentan con una variedad de laboratorios informáticos cuya misión es brindar apoyo académico a las asignaturas de las carreras de grado, como así también un servicio a los alumnos brindando un espacio para la realización de prácticas libres.
Precisamente el contexto de esta investigación es un laboratorio informático y está orientado al estudio actividades de gestión y prevención de incidentes en equipos informáticos. El mismo funciona en la Universidad Tecnológica Nacional Facultad Regional Córdoba y su actividad

depende del Departamento de Sistemas de Información.

La capacidad de cómputo del laboratorio, permite cubrir la demanda actual, ya sea la relacionada con los requerimientos de procesamiento de información de los proyectos de investigación que se están desarrollando, como así también lo solicitado por parte de las cátedras, que asisten para el dictado de clases.

El mismo dispone de siete aulas, la capacidad de máquinas de cada aula varía ya que todas no tienen el mismo tamaño, pero en promedio la capacidad de equipos es de 25 aproximadamente; y el total de equipos es 193. Las computadoras que funcionan en las aulas como en el área de Investigación y Desarrollo son computadoras personales o PC, mientras que los equipos que realizan el trabajo más importante, son computadoras con un alto grado de procesamiento que cumplen la función de servidores.

Para que las actividades académicas puedan desarrollarse con normalidad, es necesario que los recursos informáticos estén en condiciones con el objetivo de lograr la máxima disponibilidad para su normal funcionamiento. Es decir que cualquier interrupción no programada de un determinado servicio genera una disminución de la calidad de los servicios prestados.

Es fundamental que el diseño del proceso de gestión de incidentes sea óptimo, en este punto es imprescindible clarificar que es un incidente en el contexto informático. [4]
Un incidente puede considerarse como interrupciones o desmejoras de servicios que no han sido programados.
Un factor clave para evitar la ocurrencia de incidentes, está relacionado con el Mantenimiento más precisamente con el diseño de políticas de mantenimiento de carácter preventivo apropiadas para este contexto de desarrollo.

El mantenimiento moderno, se nutre de diversos enfoques y metodologías, una de ellas lo es la Ingeniería de Confiabilidad. La misma requiere un proceso cuidadoso de diagnóstico de equipos y sistemas. [5]

En este enfoque, el diagnóstico se basa en el "riesgo", puede entenderse como un proceso que busca caracterizar el estado actual y predecir el comportamiento futuro de equipos, sistemas y/o procesos, mediante el análisis del historial de fallas, datos de condición y datos técnicos, con el propósito de identificar acciones correctivas y proactivas que puedan efectivizar costos a través de la reducción de ocurrencia de fallas y eventos no deseados.
Para la estimación de la confiabilidad o probabilidad de fallas, existen dos métodos que dependen del tipo de dato disponible:

- Estimación Basada en Datos de Condición, altamente recomendable para equipos estáticos, que presentan patrones de "baja frecuencia de fallas" y por ende no se tiene un "historial de fallas" que permita algún tipo de análisis estadístico.
- Estimación Basada en el Historial de Fallas: recomendable para equipos dinámicos, los cuales por su alta frecuencia de fallas, normalmente permiten el almacenamiento de un historial de fallas que hace posible el análisis estadístico.

La confiabilidad aplicada a la ingeniería ha comprobado a lo largo de los años su eficiencia en los resultados obtenidos para la anticipación de las fallas de la operación que se encuentra en las empresas u organizaciones.
Para esta comprobación se ha tenido que desarrollar pruebas de campo aplicando la estadística.

Existen diversos estudios e investigaciones [6][7][8] que han abordado esta problemática, la gran mayoría de estudios y publicaciones están enfocados al estudio y

análisis de incidentes diversos contextos, a través de la búsqueda de un modelo adecuado para la gestión del mantenimiento.
Estas publicaciones y estudios tienen como punto en común un mismo escenario, el gran volumen de datos en el que están inmersos. Si bien la disposición de datos históricos es una fuente importante y valiosa para el estudio y análisis del comportamiento de incidentes.
Cuando el volumen de datos es significativo, se presenta una limitación a nivel de análisis, interpretación y procesamiento de los mismos.
El problema es identificar y encontrar información útil y oculta en grande bases de datos. [9]
Sería interesante disponer de un mecanismo que permita complementar las limitaciones de los métodos que ofrece la estadística descriptiva, [10] como la posibilidad de elevar los niveles de competencia de las organizaciones, basándose en la rapidez para identificar, procesar y extraer la información que realmente es importante, descubriendo conocimiento y patrones en bases de datos de volumen significativo.
Y finalmente dar respuesta a ¿Existen factores vinculados en la presentación de incidentes, con el propósito de lograr un máxima disponibilidad de los equipos informáticos?

Una alternativa atractiva para dar solución a esta problemática, es el uso de técnicas de minería de datos en el contexto de mantenimiento. Hay precedentes de que estas técnicas ya han sido utilizadas con éxito para la predicción de los incidentes, obteniendo conclusiones interesantes que demuestran que factores inciden en la presentación de los mismos.
Es importante mencionar que la mayoría de trabajos y publicaciones se dan en contexto de empresas de diferentes tipos. Del relevamiento de publicaciones e investigaciones publicadas se concluye que hay muy pocos casos de estudio relacionado con la aplicación de técnicas de minería de datos en el contexto de laboratorios informáticos.

La Minería de Datos puede definirse inicialmente como un proceso de descubrimiento de nuevas y significativas relaciones, patrones y tendencias al examinar grandes cantidades de datos. La disponibilidad de grandes volúmenes de información y el uso generalizado de herramientas informáticas ha transformado el análisis de datos orientándolo hacia determinadas técnicas especializadas englobadas bajo el nombre de Minería de Datos. [11]
Existen dos enfoques o áreas dentro del Data Mining: las técnicas clásicas de Estadística y las de minería de datos propiamente dicha que derivan, en su mayoría, de los aportes de la Inteligencia Artificial como las redes neuronales, lógica borrosa y mecanismos de aprendizaje automático. [12]
Las funciones estadísticas ofrecen diversos métodos de pronóstico para dar apoyo al proceso de toma de decisiones. Aunque no son propias de minería de datos (se utilizaban mucho antes de que surgiera este concepto), resultan de gran utilidad a la hora de descubrir patrones o elaborar modelos de predicción. Pueden emplearse para obtener más información sobre los datos, lo que permitirá tomar decisiones más acertadas cuando se apliquen los procesos de minería. Algunos ejemplos de estas técnicas son la regresión lineal, el Análisis Factorial, el Análisis de Componentes Principales entre otras.

Las técnicas de Minería de Datos persiguen el descubrimiento automático del conocimiento contenido en la información almacenada de modo ordenado en grandes base de datos. Estas técnicas tienen como objetivo descubrir patrones, perfiles y tendencias a través del análisis de los datos utilizando tecnologías de reconocimiento de patrones, redes neuronales, lógica difusa, algoritmos genéticos y otras técnicas avanzadas.

La clasificación inicial de las técnicas de minería de datos distingue entre técnicas predictivas, en las que las variables pueden clasificarse inicialmente en dependientes e independientes y técnicas descriptivas, en las que todas las variables tienen inicialmente el mismo status.

El objetivo principal de esta investigación es la detección de factores y posibles relaciones de los incidentes informáticos de hardware, ocurridos en el laboratorio de sistemas de información, mediante el uso de métodos estadísticos y técnicas de minería de datos en el periodo 2004 al 2012.
De este objetivo principal se desprenden los siguientes objetivos específicos:

- Determinar la situación actual de políticas de mantenimiento aplicadas en el laboratorio informático.

- Caracterizar la evolución y situación actual de ocurrencia de incidentes informáticos ocurridos en el laboratorio informático, mediante la selección de variables significativas.

- Determinar el grado ocurrencia y posibles relaciones involucradas en la presentación de fallas, para prevenir y disminuir la ocurrencia de incidentes informáticas.

Con la ejecución de este proyecto de investigación se pretende aplicar en primera instancia métodos de la estadística descriptiva, con el propósito de caracterizar los datos de la muestra considerada. Y posteriormente combinar el uso de métodos pertenecientes a la rama de la Minería de Datos, como las técnicas de asociación entre otras, con el propósito de obtener reglas que permitan identificar variables y posibles relaciones, que nos permitan determinar cual/es de ellas tienen mayor influencia en la presentación de incidentes informáticos.

En base a las predicciones resultantes, los algoritmos de asociación son capaces de determinar la clase de acciones de carácter preventivo.

Con la implementación y construcción de un modelo de conocimiento permitirá conocer el comportamiento de los incidentes en el periodo considerado, facilitando la elaboración de un plan de prevención logrando la disminución de los reportes de incidentes.
Con esto se logra una mayor disponibilidad de los equipos informáticos para las diversas actividades academicas que se desarrollan en el Laboratorio de Sistemas. Desde el punto de vista económico se logra disminuir los costos relacionados con la adquisición de determinados insumos o componentes/piezas que son utilizados en las tareas de mantenimiento.

Para poder dar respuesta al objetivo principal de esta investigación, se ha seleccionado dentro de las técnicas predictivas las que corresponden al modelo de clasificación y asociación.
Las técnicas de asociación, objeto de estudio en este proyecto, permite la realización de un proceso inductivo mediante el cual se lleva a cabo la generación de un conjunto de reglas de decisión con el proposito de obtener hipótesis que traten de explicar un determinado sistema.[13,14, 15] El sistema es representado por un conjunto de ejemplares y las reglas de decisión representarán conceptos que describen dicho sistema. Al modelar cada uno de los diferentes eventos que pueden ocurrir en el sistema, se esta realizando el aprendizaje inductivo de conceptos.
Con respecto a la sintaxis de las reglas, el antecedente corresponde a una condición, simple o compleja, que se ha de cumplir para que la regla se dispare y consecuentemente, se seleccione el concepto al que representa.
Se ha realizado un relevamiento investigando si existen estudios

relacionados con la aplicación de la Minería de Datos que se focalicen en el tratamiento de incidentes de los equipos en laboratorios informáticos. Actualmente no existen estudios previos relacionados con el campo del Mantenimiento.

Teniendo en cuenta los resultados esperados, los mismos podrían impactar en el ámbito del mantenimiento y la fiabilidad de equipos informáticos.
La metodología a desarrollar puede ser tomada como referencia y ser aplicada en el contexto de los laboratorios informáticos de entidades públicas, provinciales y nacionales.

Elementos del Trabajo y metodología

Esta investigación utiliza un enfoque mixto, es decir que para poder aproximarse al conocimiento combina el enfoque cuantitativo y cualitativo.
El enfoque mixto permite definir la forma de proceder en la búsqueda de conocimiento combinando y complementando metodologías y técnicas, vincula y analiza los datos cualitativos y cuantitativos para dar respuesta a la pregunta del problema central de esta investigación.
Esta forma de trabajo permite superar de las dificultades debido al uso de una sola metodología.
De acuerdo a los objetivos específicos planteados en la sección anterior, se plantea la especificación de las etapas metodológicas a cumplir.

Elaboración de un marco teórico que fundamente la investigación. Para lo cual se llevará a cabo una búsqueda de investigaciones, publicaciones y revistas especializadas accesibles por internet relacionadas con el tema; como consulta a bibliografía relevante del campo bajo estudio.

Con el fin de determinar la situación actual acerca de las políticas de mantenimiento preventivo aplicadas a los recursos informáticos en el laboratorio, se utilizará como método de recolección de datos entrevistas de carácter no estructurado al personal del área técnica. En la misma se indagarán sobre aspectos operativos de los mantenimientos y la existencia de herramientas de registración de incidentes informáticos.

Para poder caracterizar la evolución y situación actual de ocurrencia de incidentes informático, se realizará un análisis y recolección de los datos evaluando la posibilidad de seleccionar el formato adecuado para su interpretación y descripción de los incidentes.
Luego se llevará a cabo una triangulación de los datos provenientes de las entrevistas no estructuradas realizadas al personal de área técnico y de la planilla de incidente usada para el reporte de los mismos.
Seleccionado los datos significativos, se analizará el desarrollo de una herramienta adecuada que permita y facilite la carga de los datos históricos referidos incidentes.
El instrumento a usar para caracterizar los datos de los incidentes almacenados, son métodos estadísticos de carácter descriptivo e inferencial.

Los datos recolectados de los registros de incidentes pueden ser impuros. Las causas de esta situación pueden ser variadas como datos incompletos, con ruido e inconsistentes, lo que puede conducir que la extracción de reglas sean poco confiables.
Finalizada la carga de incidente se prevé en la fase de preparación de los datos la aplicación de técnicas de análisis de datos que permite mejorar la calidad de los datos, con el propósito de mejorar la eficiencia del proceso de Minería de Datos.

Las tareas que se han ejecutado son las de Recolección de Datos e Integración, y Reducción de Datos.

En la fase de recolección los datos se obtuvieron como resultado de las entrevistas realizadas al personal de área técnica como y de la planilla de registración de incidentes. Como todo el proceso de registración del incidente y el mantenimiento correctivo es realizado de forma manual, fue necesario desarrollar una aplicación para la carga de los datos y posterior procesamiento. Para poder implementar esta tarea fue necesario hacer una selección de los datos significativos para el estudio, como así también la tipificación de los datos nominales con el propósito de facilitar la etapa de transformación de datos.

Las variables selecionadas y consideras para el estudio se detallan a continuación:

- Anio reporte
  - Tipo de variable: Cuantitativa-Discreta.
  - Escala: Intervalo.
  - Valores posibles: No aplica
- Hora reporte
  - Tipo de variable: Cuantitativa-Continua.
  - Escala: Intervalo.
  - Valores posibles: No aplica
- Área
  - Tipo de variable: Cualitativa.
  - Escala: nominal.
  - Valores posibles: Encargado de turno, Encargado de aula, Área Técnica, Área de Operadores de red, Sin Datos
- Turno de reporte
  - Tipo de variable: Cualitativa.
  - Escala: Nominal.
  - Valores posibles: Mañana, Tarde, Noche, Sin Datos.
- Origen Máquina:
  - Tipo de variable: Cualitativa.
  - Escala: Nominal
  - Valores posibles: Gabinete, Servidor.
- Número Máquina:
  - Tipo de variable: Cualitativa
  - Escala: Nominal.
  - Valores posibles: No aplica.
- Reporte de Fallo:
  - Tipo de variable: Cualitativa.
  - Escala: Nominal.
  - Valores posibles: Fuentes, Coolers de Fuente, Micro, Coolers de Micro, Memorias, Discos Duros, Placa Madre, Placa de Red, Placa de Video, Otros.
- Fecha_reparación1:
  - Tipo de variable: Cuantitativa-Continua.
  - Escala: Intervalo.
  - Valores posibles: No aplica.
- Fallo_real_1:
  - Tipo de variable: Cualitativa.
  - Escala: Nominal
  - Valores posibles: Fuentes, Coolers de Fuente, Micro, Coolers de Micro, Memorias, Discos Duros, Placa Madre, Placa de Red, Placa de Video, Otros.
- Resuelto_reparación_2:
  - Tipo de Variable: Cualitativa.
  - Escala: Nominal.
  - Valores posibles: Si, No, Sin Datos.
- Fecha_reparación2:
  - Tipo de dato: Fecha.
  - Tipo de Variable: Cuantitativa-Discreta.
  - Escala: Intervalo.
  - Valores posibles: No aplica.
- Fallo_real_2:
  - Tipo de variable: Cualitativa.
  - Escala: Nominal.
  - Valores posibles: Fuentes, Coolers de Fuente, Micro, Coolers de Micro, Memorias, Discos Duros, Placa Madre, Placa de Red, Placa de Video, Otros.
- Resuelto_reparación_2:
  - Tipo de dato: Lógico
  - Tipo de variable: Cualitativa.
  - Escala: Nominal.
  - Valores posibles: Si, No, Sin Datos.

Finalizada esta etapa se realizarán una serie de pruebas con algoritmos de clasificación y asociación, con el objetivo de seleccionar aquellos que proporcionen los mejores resultados de clasificación.

En la etapa de interpretación de resultados, esta planificado la realización de una reunión entre el grupo de investigadores y el personal del área técnica del laboratorio, con el objetivo de validar las reglas de asociación con mayor grado de confiabilidad, según su experiencia.
Finalmente se procederá la elaboración y comunicación de una propuesta que incluirá una serie de recomendaciones, que permita concretar a futuro una gestión óptima de las tareas de mantenimiento logrando la disminución de incidentes.

Resultados

Si bien proyecto de investigación si bien se encuentra en su primer año de ejecución prevista, los resultados obtenidos hasta el momento, están relacionados con la caracterización de las políticas de mantenimiento ejecutadas en el laboratorio.
De la entrevista realizada por el grupo de investigador con el personal de área técnica, se pudieron conocer los siguientes aspectos.
El personal del área técnica, está conformado por graduados, alumnos y becarios cuyo cumplimiento de horario es rotativo distribuido entre la mañana, tarde y noche.
La operatoria de registración de las intervenciones o mantenimiento correctivo llevada a cabo por el personal de área técnica es el, es chequear diariamente la planilla de incidentes que es reportada por los encargados de turno o encargados de aula.
En la planilla de incidentes se consignan los siguientes datos:

Correspondientes al incidente:

- Fecha y Hora
- Responsable de reporte Incidente
- Número de Aula
- Número de Máquina
- Descripción de incidente

Referidos al mantenimiento correctivo.

- Fecha y Hora de mantenimiento
- Técnico responsable
- Descripción del mantenimiento
- Indicador si el mantenimiento está resuelto.

Cabe destacar que el procedimiento de gestión de incidentes es totalmente manual, por lo que para que el procesamiento sea posible fue necesario analizar los datos o atributos significativos para el análisis de incidente.

La población total está conformada por 700 registros de incidentes y la muestra considerada para el análisis es:
Set de datos:

- 173 registros (muestra inicial)
- 16 atributos

Para el cálculo del tamaño de la muestra se utilizó uno de los métodos a partir de una proporción, mediante la siguiente fórmula:

$$n=t^2*p\ (1-p)/e^2$$

Donde:

n=tamaño de la muestra.
t = nivel de confianza deducido a partir de la tasa de confianza (tradicionalmente 1,96 para una tasa de confianza del 95%)
p = proporción aproximada de la población que presenta la característica estudiada en el estudio. Cuando esta proporción se ignora, se puede realizar un pre-estudio o sino p = 0,5.
e = margen de error (tradicionalmente fijado en el 5%).

Para este estudio el nivel de confianza establecido fue de un 90% y el margen de error del 5%.

En la etapa de preparación de datos, fue necesario determinar [16] la estructura de los datos para poder entender este concepto es necesario definir el término conjunto de datos, este hace referencia a los datos que

serán utilizados por el modelo de minería de datos para encontrar patrones.

En este paso fue necesario efectuar procesos para el modelado de los datos como la tipificación de datos.
Los datos que se aplicó este proceso, fueron los siguientes:

- Turno
- Área
- Origen máquina
- Reporte Fallo (Diagnóstico)
- Fallo real (Fallo incidente)

Los primeros resultados que permiten caracterizar los datos de estudio se pueden visualizar en la siguiente figura.

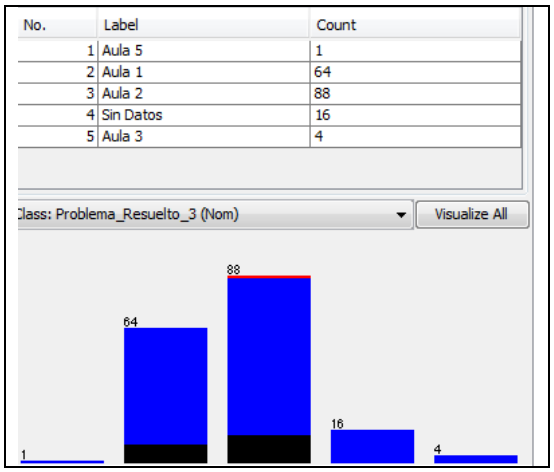

Figura 1: Atributo Aula (origen del incidente)

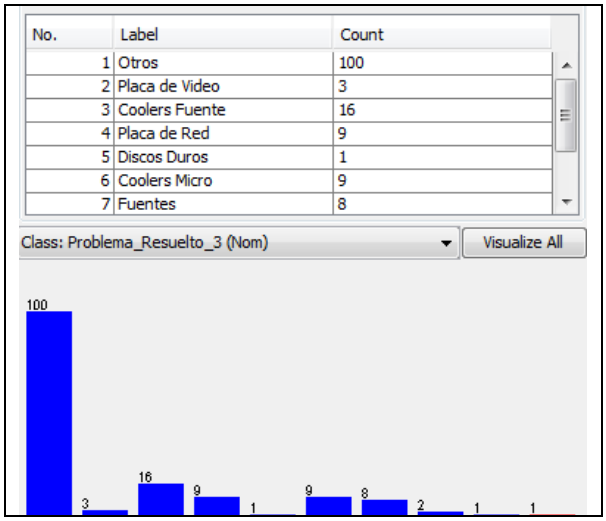

Figura 2: Atributo Fallo Real (origen del incidente)

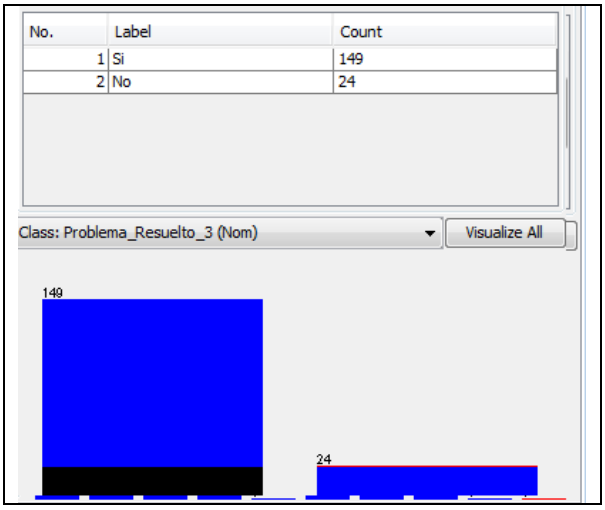

Figura 3: Reporte Fallo (causa del incidente)

Analizando los resultados arrojados en la figura 2 se puede evidenciar una alta ocurrencia de incidentes informáticos cuyo origen son "Otros". Para obtener mejores resultados, se decidió hacer un análisis de los registros de la muestra haciendo foco en este atributo, se decidió la incorporación de otros valores posibles como: Monitor, Mouse, Teclado y Lectora.

Cabe destacar que estos resultados parciales han sido obtenidos de la muestra obtenida, algunas de las inferencias preliminares que surgen de la fase exploratoria de los datos se sintetizan a continuación:

- La mayoría del reporte de incidentes fue reportado por el área de encargado de turno. La fundamentación de este aspecto es que ante la presentación de un problema o dificultad, la gran mayoría de las veces es reportada ya sea por docentes o alumnos a esta área.
- Respecto al reporte de incidentes la causa más frecuente de presentación son "Otras causas", en segundo lugar "Coolers de Fuente" y en tercer lugar "Placa de Red" y "Coolers de Micro" en igual proporción.
- La presentación de incidentes se presenta mayormente en el Aula 1 y Aula 2, que son la que tienen mayor capacidad.
- Otro aspecto que se puede concluir es que la solución de los reporte de incidentes en la primera instancia de intervención por

el personal de área técnica son reparados y el problema se logra subsanar.

Respecto a los integrantes del grupo de investigación, uno de los integrantes de esta línea de investigación está desarrollando su Tesis de Magister Ingeniería en Sistemas de Información, los demás son docentes con categoría de investigador y se desempeña un alumno becario.
Se dirigen trabajos de tesinas de la Facultad Regional Córdoba en temas relacionados con el proyecto.

Discusión

La información que se encuentra analizada tiene un valor inestimable. No obstante, esa información; no es capaz por sí sola de dar respuestas, hasta tanto no se efectúe sobre ella un trabajo de análisis e interpretación. Para entender el comportamiento y evolución hasta la actualidad será necesario obtener un modelo de conocimiento con el propósito de analizar las reglas de asociación más significativas y que permitan caracterizar y describir de manera adecuada los incidentes informáticos en el contexto del laboratorio informático.

La difusión y publicación de la metodología propuesta para la búsqueda de variables y posibles relaciones de incidentes informáticos puede ser aplicada a problemas de incidentes en laboratorios informáticos pertenecientes a distintos niveles educativos para la gestión eficiente de incidentes y tareas de mantenimiento.

Algunas líneas de investigación futuras que pueden derivar de esta investigación son:

- Desarrollo de algoritmos de minería de datos específicos para problemas relacionados con la gestión de incidentes informáticos y mantenimiento.
- Mejorar la metodología para la obtención de modelos de conocimiento de acuerdo a los requerimientos de laboratorios informáticos.
- Desarrollo de herramientas informáticas de aprendizaje automática y minería de datos más sencillas e intuitivas de utilizar, para ser utilizadas por personas que trabajan en el ámbito de mantenimiento.
- Integración de nuevos algoritmos de clasificación y asociación en herramientas de aprendizaje y minería de datos ya existentes.
- Optimizar la metodología para la selección de variables significativas e indicadores para el análisis y estudio de gestión de incidentes y mantenimiento.

Conclusión

Este trabajo presenta como alternativa de solución al problema de incidentes informáticos, la utilización de técnicas de minería de datos para lograr describir el comportamiento de los mismos y descubrir factores y posibles relaciones vinculados en la ocurrencia de los mismos.

Al concluir esta investigación se pretende lograr el descubrimiento de factores y relaciones entre variables que influyen en la presentación de incidentes en los equipos informáticos, permitirá lograr ventajas y planificar diferentes aspectos como:

- Identificación del él/los posible/s origen/es de incidentes, como por ejemplo conocer si la presentación de incidentes está relacionado mayormente por inconvenientes técnicos o por malos hábitos de operación de los equipos.
- Elaboración de procedimientos técnicos y tareas de mantenimiento a efectuar periódicamente.
- Capacitación al personal de Área Técnica.
- Disminución de los tiempos muertos o de parada de los equipos.
- Aprovechamiento y uniformidad en la carga de trabajo del personal de Área, debido a una planificación de actividades.
- Diseño de un proceso de compras de componentes o piezas de los equipos, que

facilite la disponibilidad de un stock aceptable.

Referencias

[1] Yelitza Josefina Marcano Aular y Rosalba Talavera Pereira;"Minería de Datos como soporte a la toma de decisiones";(2007;) ISSN 1012-1587
Link de consulta:
http://www.scielo.org.ve/scielo.php?pid=S1012-15872007000100008&script=sci_arttext

[2] Karla Ivette King Núñez;"Ingeniería de Confiabilidad: Antecedentes y herramientas",(2012), Link de consulta:http://www.gestiopolis.com/administracion-estrategia-2/ingenieria-de-confiabilidad.htm

[3] Abel Meriño Ibarra; "El laboratorio de informática", Link de consulta: http://www.monografias.com/trabajos14/laborat-docente/laborat-docente.shtml

[4] Br. Manuel de Jesús Gómez Torrejano, "Aplicación de la Minería de Datos para describir las ocurrencias de fallas en los equipos y sistemas de la plataforma tecnológica de PDVSA AIT Servicios Comunes Centro"

[4] Teruel Alejandro, "Gestión de Incidentes y Gestión de Problemas"; (2010), Link de consulta: https://docs.google.com/viewer?a=v&pid=sites&srcid=ZGVmYXVsdGRvbWFpbnxnZXN0aW9uc2VydmljaW9zaW5mb3JtYXRpY29zfGd4OmVhNGM3ZTBiNzA2NTI4NQ, Fecha de consulta: 16 de mayo, 20:00hs

[5] Mike Barnes; "Las 10 principales percepciones erróneas acerca del análisis de aceite";
Link de consulta:http://confiabilidad.net/articulos/pmo-optimización-de-mantenimiento/

[6] Diana Cardoso Abio, "Minería De Datos En La Detección De Intrusos", (2011), Sitio de consulta: http://www.buenastareas.com/ensayos/Miner%C3%ADa-De-Datos-En-La-Detecci%C3%B3n/3255401.html, Fecha de consulta: 13 de mayo, 16:00hs

[7] Diego Vallejo, Germán Tenelanda; "Minería de Datos aplicada en detección de intrusos", Ing. USBMed, Vol. 3, No. 1, Enero-Junio 2012

[8] Manuel de Jesús Gómez Torrejano ;"Aplicación de la Minería de Datos para describir las ocurrencias de fallas en los equipos y sistemas de la plataforma tecnológica de PDVSA AIT Servicios Comunes Centro"

[9] M. N. Quadril and N. V. Kalyankar, "Drop Out Feature of Student Data for Academic Performance Using Decision Tree Techniques", Global Journal of Computer Science and Technology, vol. 10, pp. 2-5, 2010

[10] Yelitza Josefina Marcano Aular y Rosalba Talavera Pereira;"Minería de Datos como soporte a la toma de decisiones";(2007;) ISSN 1012-1587
Link de consulta:
http://www.scielo.org.ve/scielo.php?pid=S1012-15872007000100008&script=sci_arttext

[11] Pérez López Cesar, Santín González Daniel;"Minería de Datos. Técnicas y Herramientas", Editorial: Thomson (2007), ISBN:

[12] Lucena, María;"Proyecto de Fin de Carrera: Análisis Comparativo entre Métodos Estadísticos y de Minería de Datos", (2005), http://www.iit.upcomillas.es/pfc/resumenes/42b974d2de20c.pdf, Fecha de consulta: 14 de mayo, 08:00hs

[13] Pajares Martín-Sanz, De la Cruz García Jesús; Aprendizaje Automático. Un enfoque práctico, 2010, ISBN: 9788499640112

[14] Basilio Sierra Araujo, Aprendizaje Automático: conceptos básicos y avanzados. Aspectos prácticos utilizando el software Weka.ISBN 10: 84-8322-318-X, 2006

[15] Orallo Hernandez José, Ramírez José, Ferri Cesar; Introducción a la Minería de Datos, 2004, ISBN: 8420540919

[16] Merlino Hernán, "Metodología de transformación de datos para su explotación", Página web consultada:
http://laboratorios.fi.uba.ar/lsi/R-rtis-6-2-2004.pdf,
Fecha de consulta: 11 de Mayo, 15:00hs.

[17] T. Menzies, Y. Hu. Data Mining For Busy People. IEEE Computer, (2003), pgs. 18-25.

[18] Cuadras M Carles, "Nuevos Métodos de Análisis Multivariante", CMC Editions, Barcelona España,(2012),http://www.ub.edu/stat/personal/cuadras/metodos.pdf

[19] Nickhil R. Pal, Lakhmi Jain;Advanced Techniques in Data Mining and Knowledge Discovery,Springer,2005,ISBN:1-85233-867-9

Datos de Contacto:

Cynthia Lorena Corso. Universidad Tecnológica Nacional. Maestro M. López esq. Cruz Roja Argentina Ciudad Universitaria.
E-mail:cynthia@bbs.frc.utn.edu.ar